January 1989

Report No. STAN-CS-89-1244

3

AD-A204 526

Software Performance on Nonlinear Least-Squares Problems

by

Christina Fraley

Department of Computer Science

Stanford University

Stanford, California 94305

89 2 15 254

CLaSSiC Project
Report CLaSSiC-88-06

June 1988

Software Performance on Nonlinear Least-Squares Problems

Christina Fraley

Center for Large Scale Scientific Computation
Building 460, Room 313
Stanford University
Stanford, California 94305

(A)

Software Performance on Nonlinear Least-Squares Problems

Christina Fraley
Stanford University†

Abstract

This paper presents numerical results for a large and varied set of problems using software that is widely available and has undergone extensive testing. The algorithms implemented in this software include Newton-based linesearch and trust-region methods for unconstrained optimization, as well as Gauss-Newton, Levenberg-Marquardt, and special quasi-Newton methods for nonlinear least squares. Rather than give a critical assessment of the software itself, our original purpose was to use the best available software to compare the underlying algorithms, to identify classes of problems for each method on which the performance is either very good or very poor, and to provide benchmarks for future work in nonlinear least squares and unconstrained optimization. The variability in the results made it impossible to meet either of the first two goals; however the results are significant as a step toward explaining why these aims are so difficult to accomplish. (KR)

† present address: Université de Genève COMIN

Acknowledgements

I am fortunate to have had generous financial support for this research from the following sources:

- the Xerox Corporation, in the form of a fellowship,
- Joseph Oliger, under Office of Naval Research contract # N00014–82–K–0335,
- Jean-Philippe Vial, under Fonds National de la Recherche Scientifique Suisse research grant # 1 467–0.86,
- the Systems Optimization Laboratory, Stanford University, under National Science Foundation grant # CCR–8413211.

I am also indebted to Stanford Linear Accelerator Center for the use of their computer facilities.

1. Introduction

The nonlinear least-squares problem is that of minimizing a sum of squares

$$\min_{x \in \Re^n} \frac{1}{2} \sum_{i=1}^{m} \phi_i(x)^2,$$

in which each ϕ_i is a real-valued function having continuous second partial derivatives. The problem can also be posed as a minimization of the l_2 norm of a multivariate function:

$$\min_{x \in \Re^n} \frac{1}{2} \|f(x)\|_2^2,$$

where

$$f(x) = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \\ \vdots \\ \phi_m(x) \end{pmatrix}.$$

We shall refer to the function $\frac{1}{2}\|f(x)\|_2^2$ as the nonlinear *least-squares objective* function. It is assumed that n and m are relatively small, so that algorithms are not formulated with any special considerations for limited storage. A common instance is the choice of parameters β within a nonlinear model φ:

$$\min_{\beta \in \Re^n} \sum_{i=1}^{m} \frac{1}{2} (\varphi(\beta; \tau_i) - d_i)^2,$$

where d_i are observations at prescribed values τ_i.

Many specialized algorithms have been developed to take advantage of the structure of the nonlinear least-squares objective. This paper presents numerical results for a large and varied set of problems using software that is widely available and has undergone extensive testing. The algorithms implemented in this software include Newton-based linesearch and trust-region methods for unconstrained optimization, as well as Gauss-Newton, Levenberg-Marquardt, and special quasi-Newton methods for nonlinear least squares. Rather than give a critical assessment of the software itself, our original purpose was to use the best available software to compare the underlying algorithms, to identify classes of problems on which the performance of each method is either very good or very poor, and to provide benchmarks for future work in nonlinear least squares and unconstrained optimization. The variability in the results made it impossible to meet either of the first two goals. However, the results are significant in that they reveal a great deal about the reasons these aims why aims are so difficult to accomplish. For treatment of issues and

methodology for software performance evaluation see, e. g., Moré, Garbow, and Hillstrom [1978; 1981], Hiebert [1979; 1981], and Hanson and Krogh [1987]. Hiebert [1979; 1981] conducts an extensive evaluation of twelve programs for nonlinear least squares, in which she includes software that uses first derivatives as well as some that does not. In the present study, all of the software requires first if not second derivatives of the problem functions.

This paper is organized as follows. Section 2 reviews computational techniques for the unconstrained optimization problem. These methods are of interest because the nonlinear least squares problem is a particular instance of unconstrained optimization, so that special-purpose algorithms for sums of squares should compare favorably in performance with those developed for the more general case. Moreover, much of the motivation for unconstrained optimization methods is also relevant to algorithm development for nonlinear least squares. Our emphasis is on computational issues related to the methods included in this study. For more extensive treatment of unconstrained optimization algorithms, see Fletcher [1980], Gill, Murray, and Wright [1981], Dennis and Schnabel [1983], and Moré and Sorensen [1984]. Section 3 reviews the principal approaches that are used in software for nonlinear least-squares problems. These are Gauss-Newton methods; Levenberg-Marquardt methods, one of which is implemented in the software package MINPACK [Moré (1978), Moré, Garbow, and Hillstrom (1980)]; corrected Gauss-Newton methods [Gill and Murray (1978)], which are the basis for the NAG Library nonlinear least-squares software; and methods that form quasi-Newton approximations to the term $B = \sum_{i=1}^{m} \phi_i \nabla^2 \phi_i$ in the nonlinear least-squares Hessian, a strategy that is adaptively combined with a Gauss-Newton method and a Levenberg-Marquardt method in the computer algorithm NL2SOL [Dennis, Gay, and Welsch (1981a, 1981b)]. We assume a knowledge of numerical techniques for linear least-squares (e. g., Lawson and Hanson [1974], and Golub and Van Loan [1983]). For more information on algorithms specific to nonlinear least-squares problems, see Fraley [1988] and the references cited in that paper. Section 4 is a summary and discussion of the numerical results. Section 5 contains tables of all of the results, as well as information about the software and test problems used in obtaining them.

1.1 Definitions and Notation

We shall use the following definitions and notational conventions :

- Generally subscripts on a function mean that the function is evaluated at the corresponding subscripted variable (for example, $f_k = f(x_k)$). An exception is made for the residual functions ϕ_i, where the subscript is the component index for the vector f.

- f - The vector of nonlinear functions whose l_2 norm is to be minimized.
The nonlinear least-squares problem is
$$\min_{x \in \Re^n} \frac{1}{2} f(x)^{\mathrm{T}} f(x),$$
where the factor $\frac{1}{2}$ is introduced in order to avoid a factor of two in the derivatives.

- ϕ_i - The ith *residual* function, also the ith component of the vector f.
$$f(x) \equiv \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_m(x) \end{pmatrix}.$$
An alternative formulation of the nonlinear least-squares problem is
$$\min_{x \in \Re^n} \frac{1}{2} \sum_{i=1}^{m} \phi_i(x)^2,$$
where each $\phi_i(x)$ is a smooth function mapping $\Re^n$ to $\Re$.

- J - The Jacobian matrix of f.
$$J(x) \equiv \nabla f(x) = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_m}{\partial x_1} & \cdots & \frac{\partial \phi_m}{\partial x_n} \end{pmatrix}$$

- g - The gradient of the nonlinear least-squares objective.
$$g(x) \equiv \nabla \left(\frac{1}{2} f(x)^{\mathrm{T}} f(x) \right) = J(x)^{\mathrm{T}} f(x)$$

- B - The part of the Hessian matrix of the nonlinear least-squares objective that involves second derivatives of the residual functions.
$$\nabla^2 \left(\frac{1}{2} f(x)^{\mathrm{T}} f(x) \right) = J(x)^{\mathrm{T}} J(x) + B(x),$$

where

$$B(x) \equiv \sum_{i=1}^{m} \phi_i(x) \nabla^2 \phi_i(x).$$

- $\mathcal{R}(A)$ - The *range* of A.

  If A is an $m \times n$ matrix, then

  $$\mathcal{R}(A) \equiv \{b \mid Ax = b \text{ for some } x \in \Re^n\}$$

  is a subspace of $\Re^m$.

- $\mathcal{N}(A)$ - The *null* space of A.

  If A is an $m \times n$ matrix, then

  $$\mathcal{N}(A) \equiv \{z \mid Az = 0\}$$

  is a subspace of $\Re^n$. $\mathcal{N}(A)$ is the orthogonal complement of $\mathcal{R}(A^T)$ in $\Re^n$.

2. Methods for Unconstrained Optimization

2.1 Local Quadratic Approximation

Software for the unconstrained optimization problem

$$\min_{x \in \Re^n} \mathcal{F}(x).$$

is usually based on successive minimization of a quadratic approximation

$$Q_k(p) \equiv \nabla \mathcal{F}_k^T p + \frac{1}{2} p^T H_k p \qquad (2.1.1)$$

for $\mathcal{F}(x_k + p) - \mathcal{F}(x_k)$, the change in $\mathcal{F}$ at x_k. The matrix H_k is positive definite, so that the model $Q_k(p)$ has a well-defined minimum

$$p_k = -H_k^{-1} \nabla \mathcal{F}_k,$$

that can be computed efficiently. Positive definiteness of H_k also means that p_k is a descent direction for $\mathcal{F}$ at x_k, which is essential for linesearch methods (see Section 2.2.1). For methods based on (2.1.1), the condition

$$\lim_{k \to \infty} \frac{\left\| (H_k - \nabla^2 \mathcal{F}(x^*)) p_k \right\|}{\| p_k \|} = 0 \qquad (2.1.2)$$

is equivalent to local superlinear convergence of the sequence $\{x_k + p_k\}$ to an isolated local minimum x^* of $\mathcal{F}$ (see Dennis and Moré [1974; 1977]). The relationship (2.1.2) implies that the step p_k approaches the Newton step in both magnitude and direction, although the sequence of matrices $\{H_k\}$ need not converge to $\nabla^2 \mathcal{F}(x^*)$.

Section 2.2 is concerned with modifications that are used to enforce convergence from an arbitrary starting point. These modifications fall into two categories: *linesearch* methods and *trust-region* methods. Section 2.3 deals with the choice of H_k in (2.1.1), so that condition (2.1.2) for superlinear convergence is satisfied. We discuss methods that use exact second derivatives as well as quasi-Newton approximations.

2.2 Globalization Strategies

2.2.1 Linesearch Approach

Linesearch methods obtain a new iterate in two essentially separate phases. First, a *descent direction* p_k is found for $\mathcal{F}$; that is, a vector p_k is computed for which

$$\nabla \mathcal{F}_k^T p_k < 0. \tag{2.2.1}$$

Condition (2.2.1) is equivalent to saying that $\mathcal{F}$ initially decreases along the direction p_k from x_k. Various ways of defining p_k are discussed in Section 2.3. This section is concerned with the second phase of a linesearch method, that of finding a steplength α_k satisfying

$$\mathcal{F}(x_k + \alpha_k p_k) < \mathcal{F}(x_k), \tag{2.2.2}$$

once a descent direction is obtained.

Because of (2.2.1), condition (2.2.2) can be satisfied by choosing a sufficiently small value of α_k, but the result may not be an appreciable reduction in $\mathcal{F}$. In fact, $\{x_k + \alpha_k p_k\}$ may converge to a point that is not a stationary point unless conditions stronger than (2.2.2) are imposed on α_k (see, e. g., Dennis and Schnabel [1983], Chapter 6). On the other hand, finding a minimum of $\mathcal{F}$ along p_k is an iterative process which could require many function and derivative evaluations. Steplength algorithms instead compute α_k that satisfies conditions sufficient to ensure convergence to a *stationary point* of $\mathcal{F}$ whenever the sequence $\{p_k\}$ is bounded away from orthogonality to the gradient.

The work of Goldstein [1965; 1967], Armijo [1966], Goldstein and Price [1967], and Wolfe [1969; 1971], (see also Ortega and Rheinboldt [1970]) established the fundamental principles underlying most steplength algorithms. A simple strategy for sufficient decrease is based on the condition

$$\mathcal{F}(x_k + \alpha_k p_k) - \mathcal{F}(x_k) \leq \mu \alpha_k \nabla \mathcal{F}_k^T p_k, \tag{2.2.3}$$

for $\mu \in [0, 1)$. An initial value (usually $\alpha_k = 1$) is tried first, and then a backtracking strategy is used to reduce it until an admissible step is found. The steplength strategy of Gill et al. [1979], combines (2.2.3) with the condition

$$|\nabla \mathcal{F}(x_k + \alpha_k p_k)^T p_k| \leq -\eta \nabla \mathcal{F}_k^T p_k, \tag{2.2.4}$$

for $\eta \in [0, 1)$, which keeps the steplength bounded away from zero by forcing it to approximate a local minimum of $\mathcal{F}$ along p_k. A procedure for one-dimensional minimization is truncated, using

(2.2.4) as the criterion for termination. This is accomplished by polynomial interpolation to the function

$$\Phi(\alpha) \equiv \mathcal{F}(x_k + \alpha p_k), \tag{2.2.5}$$

together with some safeguards to prevent iterates from being either too close together or too far apart. An exact minimization would be carried out for $\eta = 0$ in (2.2.4), while larger values of η increasingly relax this requirement. When $\mu < \eta$, an interval of steplengths satisfies both (2.2.3) and (2.2.4); if μ is chosen sufficiently small, then (2.2.3) almost always holds when (2.2.4) does. When $\mu \geq \eta$, a backtracking strategy may be used if (2.2.3) fails to hold for the steplength computed in the one-dimensional minimization. If $\nabla \mathcal{F}_k^T p_k < 0$ and α_k satisfies (2.2.3) and (2.2.4), then

$$\lim_{k \to \infty} \frac{\nabla \mathcal{F}_k^T p_k}{\|p_k\|_2} = 0,$$

which implies convergence to a stationary point of $\mathcal{F}$ provided $\{p_k\}$ remains uniformly bounded away from orthogonality to $\{\nabla \mathcal{F}_k\}$. If $\mu \leq 0.5$, both conditions (2.2.3) and (2.2.4) are automatically satisfied by superlinearly or quadratically convergent algorithms with $\alpha_k = 1$ when x_k is sufficiently close to a local minimum.

Although the theory allows considerable flexibility in choosing the interpolant to $\Phi(\alpha)$ and other parameters in the univariate minimization, as well as in the choice of μ and η in (2.2.3) and (2.2.4), in practice performance on difficult problems may be very sensitive to these factors. Moreover, safeguarding in univariate minimization requires specification of a finite interval of uncertainty in which the minimum is presumed to lie. If p_k is very large, it could happen that no satisfactory approximation to a minimum along that direction can be found, resulting in an excessively small step.

2.2.2 Trust-Region Approach

Trust-region methods were first developed for nonlinear least squares [Levenberg (1944); Morrison (1960); Marquardt (1963)] (see Section 3.2), and later independently for general unconstrained minimization [Goldfeld, Quandt, and Trotter (1966)]. Motivation for trust-region methods comes from the following observation: if the step to the unconstrained minimum of the current local model for $\mathcal{F}(x + p) - \mathcal{F}(x)$ is relatively large, then it probably falls outside the region in which the model is applicable. The basic idea is to define a neighborhood of the current

point over which an approximate minimization of a local model of the change in $\mathcal{F}$ will yield a suitable step to the next iterate.

The local model and constraints defining the neighborhood are chosen so that the subproblem has a well-defined minimum, and so that fast local convergence is possible with the unconstrained model. Typically, the model at x_k is a quadratic function $\nabla\mathcal{F}_k^T p + \frac{1}{2} p^T H_k p$, and an upper bound is imposed on a scaled l_2 norm of p, giving the subproblem

$$\min_{p\in\Re^n} \nabla\mathcal{F}_k^T p + \frac{1}{2} p^T H_k p \tag{2.2.6}$$

$$\text{subject to} \|D_k p\|_2 \le \delta_k .$$

For practical reasons, the scaling matrices D_k are usually diagonal (with positive diagonal entries). Solving (2.2.6) is equivalent to minimizing the quadratic function

$$\nabla\mathcal{F}_k^T p + \frac{1}{2} p^T \left(H_k + \lambda_k D_k^T D_k\right) p \tag{2.2.7}$$

for some $\lambda_k \ge 0$, where the matrix $H_k + \lambda_k D_k^T D_k$ is at least positive semi-definite.

In practice, it has been found to be more satisfactory to control the value of δ_k directly rather than λ_k (see Moré [1983]). Increases and decreases in δ_k are usually based on comparing the actual reduction

$$\mathcal{F}(x_k + p_k) - \mathcal{F}(x_k)$$

to the reduction predicted by the model,

$$\nabla\mathcal{F}_k^T p_k + \frac{1}{2} p_k^T H_k p_k .$$

The updating procedure for δ_k can be as simple as multiplying the current value by some prescribed factor, without compromising global convergence properties (see below). The preferred strategy for decreasing δ_k is more complicated. An approximate minimum τ_k of $\mathcal{F}(x_k + \tau p_k)$ is computed by safeguarded polynomial interpolation (as in linesearch methods — see Section 2.2.1), and $\tau_k\|D_k p_k\|_2$ is taken to be the new value of δ_k. It may be necessary to decrease δ_k a number of times before a suitable reduction in $\mathcal{F}$ is achieved and the step to a new iterate can be taken.

Once δ_k is assigned a value, it remains to find p_k when the solution to (2.2.6) is not an unconstrained minimum. Moré and Sorensen [1983] obtain λ_k in (2.2.7) by truncating a numerical procedure for computing a zero of the function

$$\Psi(\lambda) \equiv \|p_k(\lambda)\|_2 - \delta_k \equiv \left\| \left(H_k + \lambda D_k^T D_k\right)^{-1} \nabla\mathcal{F}_k \right\|_2 - \delta_k . \tag{2.2.8}$$

based on the work of Hebden [1973] (see also Reinsch [1971] and Gay [1981]). The algorithm of Gay [1983], implemented in the PORT Library [1984], approximates $p_k(\lambda)$ by a linear combination of the (scaled) steepest descent direction and the Newton direction. This technique was devised by Powell [1970] (see also Dennis and Mei [1979]), and is used because it achieves comparable performance to methods that attempt to approximate $\Psi(\lambda)$ closely, with considerably less computational effort.

Somewhat stronger convergence results have been proven for trust-region methods than are known for linesearch methods (see Section 2.2.1). Trust-region methods converge to an isolated local minimum under fairly mild conditions when exact second derivatives are used, and otherwise to a stationary point. Although global convergence properties are not affected, in practice the choice of δ_0 and the updating strategy for δ_k are important. As δ_k, and hence the norm of p, is made to approach zero, the minimizer of the quadratic becomes parallel to the steepest descent direction, $-\nabla \mathcal{F}_k$. Small values of δ_k are accordingly safe, in the sense that they guarantee a decrease, but progress may be unacceptably slow if no provision is made for taking larger steps where possible.

For more detail and general discussion of trust-region methods, see Moré [1983], Shultz, Schnabel, and Byrd [1985], and Bulteau and Vial [1987]. Related variants are described in Bulteau and Vial [1985] and Byrd, Schnabel, and Shultz [1988].

2.2.3 Stationary Points and Directions of Negative Curvature

It is possible to decrease $\mathcal{F}$ at a stationary point x^* if the Hessian matrix has one or more negative eigenvalues. The decrease is obtained by moving along a direction of negative curvature; in other words, a direction p for which $p^T \nabla^2 \mathcal{F}(x^*) p < 0$. Trust-region methods that use the quadratic model with exact Hessian information (see Section 2.3.1) will yield directions of negative curvature at stationary points when $\nabla^2 \mathcal{F}(x^*)$ is indefinite, whereas the linesearch methods discussed above terminate when the gradient vanishes.

A fundamental problem is that of deciding the length of the step to be taken along a direction of negative curvature. This problem is very much related to the problem of setting a maximum step length in order to safeguard a linesearch method, or that of determining the step bound in a trust-region method. First- and second-order information about the function at x^* indicates that an infinite step can be taken, since the quadratic part of the Taylor series at x^* is unbounded

below when $\nabla^2\mathcal{F}(x^*)$ is indefinite. Clearly an infinite step is not possible if $\mathcal{F}$ has a finite minimum.

Neither the question of choosing a direction of negative curvature, nor the problem of choosing a steplength along such directions has been adequately resolved, and thus in most methods directions of negative curvature are not explicitly sought at arbitrary points. For research on generating directions of negative curvature, and on their use in unconstrained optimization algorithms, see Gill and Murray [1974a], Fletcher and Freeman [1977], McCormick [1977], Moré and Sorensen [1979], Goldfarb [1980], and Shultz, Schnabel, and Byrd [1985].

2.3 Defining the Quadratic Model

2.3.1 Second-Derivative Methods

There are two basic frameworks for defining H_k in the quadratic model (2.1.1) when second derivative information is available: direct modification of the Hessian, and trust-region methods. Both can be viewed as procedures for producing a positive-definite quadratic model by modifying the exact Hessian $\nabla^2\mathcal{F}_k$. A method that combines the two approaches is given in Chapter 5 (Section 5) of Dennis and Schnabel [1983].

The modified Newton method of Gill and Murray [1974a] is a linesearch method in which the definition of the search direction is based on the fact that if H_k is positive definite, it can be characterized by its *Cholesky factorization*

$$H_k = R_k^T R_k, \qquad (2.3.1)$$

where R_k is upper-triangular and nonsingular (see, e. g., Stewart [1973], Chapter 3). Gill and Murray alter the Cholesky factorization procedure so that it can be continued in the event of indefiniteness or near-singularity. The modified factorization is applied to the Hessian matrix $\nabla^2\mathcal{F}_k$, resulting in the Cholesky factorization of a matrix H_k with a prescribed upper bound on its condition number. The matrix H_k may differ from $\nabla^2\mathcal{F}_k$ only in the diagonal elements. Local convergence properties of of Newton's method are preserved, because $H_k = \nabla^2\mathcal{F}_k$ whenever $\nabla^2\mathcal{F}_k$ is sufficiently positive definite. An implementation is available in the NAG Library [1984] (subroutine E04LBF). For information on other direct modification methods, see Gill, Murray, and Wright [1981], Chapter 4, and Higham [1986].

In trust-region methods with exact Hessian information, a subproblem of the form

$$\min \nabla\mathcal{F}_k^T p + \frac{1}{2} p^T \nabla^2 \mathcal{F}_k p \tag{2.3.2}$$

$$\text{subject to} \|D_k p\|_2 \le \delta_k.$$

is solved for the step p_k to the next iterate. We recall from the overview of trust-regions in Section 2.2.2 that solving (2.3.2) is equivalent to solving

$$\min \nabla\mathcal{F}_k^T p + \frac{1}{2} p^T \left(\nabla^2 \mathcal{F}_k + \lambda_k D_k^T D_k\right) p \tag{2.3.3}$$

for some non-negative value of λ_k, with $\nabla^2\mathcal{F}_k + \lambda_k D_k^T D_k$ positive semidefinite. In particular, λ_k will be positive whenever $\nabla^2\mathcal{F}_k$ is indefinite, and also when $\nabla^2\mathcal{F}_k$ is positive-definite if δ_k happens to be smaller than the norm of the scaled unconstrained minimum of the quadratic objective. In contrast to the modified Newton method described above, all of the eigenvalues of $\nabla^2\mathcal{F}_k$ are changed when $\lambda_k > 0$ in (2.3.3). As long as the constraint in (2.3.2) is inactive near a local minimum, the local convergence behavior of Newton's method is preserved. A recent implementation of a trust-region method that uses second derivatives is available in the PORT Library [1984] (subroutine DMNH; see also Gay [1983]). For further information on trust-regions with exact Hessian information, see Fletcher [1980], Chapter 5, Gay [1981], Sorensen [1982], Moré [1983], Moré and Sorensen [1983], and Shultz, Schnabel, and Byrd [1985].

2.3.2 Quasi-Newton Methods

In *quasi-Newton* methods (also called *variable metric* or *secant* methods), a sequence of approximations $H_0, H_1, \ldots,$ to the Hessian matrix of $\mathcal{F}$ is generated, with H_{k+1} depending on H_k as well as on gradient information at the current iterate. The approximate Hessian is required to satisfy the *quasi-Newton* condition

$$H_{k+1} s_k = y_k. \tag{2.3.4}$$

$$s_k \equiv x_{k+1} - x_k; \quad y_k \equiv \nabla\mathcal{F}_{k+1} - \nabla\mathcal{F}_k.$$

The quantity $y_k^T s_k$ approximates the curvature, $s_k^T \nabla^2 \mathcal{F}_k s_k$, of $\mathcal{F}$ along s_k. Equation (2.3.4) does not uniquely define H_{k+1}; papers that discuss completion of the specification include Dennis and Moré [1977], Nazareth [1984], Todd [1984], and Flachs [1986]. Conditions imposed on the approximate Hessian almost always include symmetry and positive definiteness.

It is generally agreed that the best overall performance is achieved by the BFGS update

$$H_{k+1} = H_k - \frac{H_k s_k (H_k s_k)^T}{s_k^T H_k s_k} + \frac{y_k y_k^T}{y_k^T s_k},$$

although precise reasons for its superiority are still not known (see, e. g., Brodlie [1977]). Like most proposed updates, the BFGS update is a rank-two modification of the current approximate Hessian. The BFGS update preserves positive definiteness whenever $y_k^T s_k > 0$, a condition that holds automatically in a linesearch method satisfying (2.2.4).

Originally, quasi-Newton updates were formulated in terms of H_k^{-1} rather than H_k, so that minimizing the quadratic (2.1.1) at each stage in a linesearch algorithm involved only a matrix multiplication ($\mathcal{O}(n^2)$ arithmetic operations) rather than solution of a linear system ($\mathcal{O}(n^3)$ arithmetic operations). Gill and Murray [1972] showed that rank-two quasi-Newton methods could be implemented in $\mathcal{O}(n^2)$ operations per iteration by applying an update to a Cholesky factor (see Section 2.3.1) of H_k. This has the additional advantage that it allows the numerical positive definiteness of H_k to be monitored from iteration to iteration. For more information on computational aspects of the update, see Dennis and Schnabel [1983], Chapter 9, and Gill et al. [1985].

The BFGS method belongs to a class of quasi-Newton methods that can be derived by minimizing the difference $(H_{k+1} - H_k)$ or $(H_{k+1}^{-1} - H_k^{-1})$, in various weighted norms, subject to (2.3.4) [Dennis and Schnabel (1979)]. Other classes of methods attempt to minimize the condition number of H_k by selecting parameters in a class of updates at each step [Shanno and Kettler (1970); Oren (1973, 1982); Davidon (1975); Oren and Spedicato (1976); Spedicato (1976); Schnabel (1978)]. Al-Baali and Fletcher [1985] apply a scaling factor before updating that minimizes an approximate measure of the error in the inverse Hessian matrix. Performance tests indicate that these modified methods are not as successful as the BFGS method for general problems [Brodlie (1977); Shanno and Phua (1978b); Al-Baali and Fletcher (1985)].

Under the same assumptions as required for local quadratic convergence of Newton's method, quasi-Newton methods are locally superlinearly convergent, provided H_0 is sufficiently close to $\nabla^2 \mathcal{F}(x_0)$ [Broyden, Dennis, and Moré (1973)]. Selection of the initial Hessian approximation H_0 can be critical in the performance of a quasi-Newton method. Often the identity is chosen because it gives the steepest-descent direction on the first iteration, and it is positive definite. Computational tests have shown that improved performance can sometimes be achieved by scaling H_0 before performing the first update [Shanno and Phua (1978a); Dennis and Schnabel (1983),

Chapter 9]. Another possibility is to use a finite-difference approximation to $\nabla^2\mathcal{F}(x_0)$ for H_0, modified if necessary to ensure positive definiteness. Although the choice of H_0 can have a significant effect on performance, the question of how best to choose H_0 is still open. It is generally agreed that exact or approximate curvature information should be used to start the algorithm if it is available at a reasonable cost. For a nonlinear least-squares problem, $J_0^T J_0$ can be used as the initial estimate, provided the columns of J_0 are linearly independent.

3. Methods for Nonlinear Least Squares

3.1 Gauss-Newton Methods

The Gauss-Newton method is a linesearch method in which the search direction at the current iterate minimizes the quadratic function

$$g^T p + \frac{1}{2} p^T J^T J p. \tag{3.1.1}$$

As a model for the change in the least-squares objective, (3.1.1) has the advantage that it involves only first derivatives of the residuals, and that $J^T J$ is always at least positive semi-definite. If

$$p^{GN} = \arg \min_{p \in \Re^n} g^T p \frac{1}{2} p^T J^T J p,$$

then

$$J^T J p^{GN} = -g, \tag{3.1.2}$$

so that p^{GN} is a direction of descent for $f^T f$ whenever $g \neq 0$, as required in a linesearch method. To guarantee convergence, the sequence of search directions must also be bounded away from orthogonality to the gradient, a condition that may not be met by successive Gauss-Newton directions unless the eigenvalues of $J^T J$ are bounded away from zero. Powell [1970a] gives an example of convergence of a Gauss-Newton method with exact linesearch to a non-stationary point.

The Gauss-Newton method can be viewed as a modification of Newton's method in which $J^T J$ is used to approximate the Hessian matrix

$$J^T J + \sum_{i=1}^{m} \phi_i \nabla^2 \phi_i = J^T J + B$$

of the nonlinear least-squares objective function. The assumption is that the matrix $J^T J$ should be a good approximation to the full Hessian when the residuals are small. In fact, if $f(x^*) = 0$ and $J(x^*)^T J(x^*)$ is positive definite, then the sequence $\{x_k + p_k^{GN}\}$ is locally quadratically convergent to x^*, because $J(x_k)^T J(x_k)$ is an $O(\|x_k - x^*\|)$ approximation to the Hessian of the nonlinear least-squares objective at x^*.

When J^TJ is singular, or, equivalently, when J has linearly dependent columns, (3.1.1) does not have a unique minimizer. The set of vectors that minimize (3.1.1) is the same as the set of solutions to the linear least-squares problem

$$\min_{p \in \Re^n} \|Jp + f\|_2 . \tag{3.1.3}$$

One (theoretically) well-defined alternative that is often approximated computationally is to require the unique solution of minimum l_2 norm:

$$\min_{p \in S} \|p\|_2. \tag{3.1.4}$$

where S is the set of solutions to (3.1.3), while another is to replace J in (3.1.3) by a maximal linearly independent subset of its columns. In finite-precision arithmetic, there is often some ambiguity about how to formulate and solve an alternative to (3.1.3) when the columns of J are "nearly" linearly dependent, which is significant because the numerical solution of these problems is dependent on the criteria used to estimate the rank of J. Fraley [1987b] gives some detailed examples that illustrate some of the difficulties that arise in implementation.

3.2 Levenberg-Marquardt Methods

Levenberg-Marquardt methods alter the Gauss-Newton search direction in the range of J, by replacing J^TJ in the quadratic model function with $J^TJ + \lambda D^TD$, where $\lambda \geq 0$ and D is a diagonal scaling matrix with positive diagonal entries. The step p between successive iterates minimizes the quadratic model

$$g^Tp + \frac{1}{2}p^T(J^TJ + \lambda D^TD)p, \tag{3.2.1}$$

for some $\lambda \geq 0$. Since the matrix $J^TJ + \lambda D^TD$ is positive semidefinite, minimizers p_λ of (3.2.1) satisfy the equations

$$(J^TJ + \lambda D^TD)p = -g = -J^Tf, \tag{3.2.2}$$

which are the normal equations for the linear least-squares problem

$$\min_{p \in \Re^n} \left\| \begin{pmatrix} J \\ \sqrt{\lambda}D \end{pmatrix} p - \begin{pmatrix} f \\ 0 \end{pmatrix} \right\|_2^2 . \tag{3.2.3}$$

Equivalently, p solves

$$\min_{p \in \Re^n} g^Tp + \frac{1}{2}p^TJ^TJp \tag{3.2.4}$$

$$\text{subject to } \|Dp\|_2 \leq \delta,$$

for some $\delta > 0$; that is, the Gauss-Newton quadratic model is minimized subject to a trust-region constraint.

Considerable research effort has been directed toward improvements in this class of methods since their introduction by Levenberg [1944], Morrison [1960], and Marquardt [1963] for nonlinear least-squares problems, and independently by Goldfeld, Quandt, and Trotter [1966] for general unconstrained optimization. Moré [1978] gives an implementation in which he adjusts the step bound δ in (3.2.4) rather than λ, a strategy used in trust-region methods for unconstrained optimization (see Moré [1983] for a survey). Changes in δ depend on agreement between the actual reduction in the sum of squares

$$\frac{1}{2}\left(\|f(x + p_\lambda)\|_2^2 - \|f(x)\|_2^2\right),$$

with the reduction

$$g^T p_\lambda + \frac{1}{2} p_\lambda^T J^T J p_\lambda$$

predicted by the model $J^T J + \lambda D^T D$, which is the optimum value of the objective in (3.2.4). Increases are accomplished by taking $\delta_{k+1} = 2\|D_k p_k\|_2$, while δ is decreased by multiplying by the factor $\gamma < 1$. In order to obtain λ when the bound in (3.2.4) is active, the nonlinear equation

$$\Psi(\lambda) = \|Dp_\lambda\|_2 - \delta = \left\|\left(J^T J + \lambda D^T D\right)^{-1} g\right\|_2 - \delta = 0 \tag{3.2.5}$$

is approximately solved by truncating a safeguarded Newton method based on the work of Hebden [1973] (see also Reinsch [1971]). Moré reports that, on the average, (3.2.5) is solved fewer than two times per iteration. He also proves global convergence to a stationary point of $f^T f$, without assuming boundedness for $\{\lambda_k\}$. Many computational details are given, including an efficient method for calculating the derivative of $\Psi(\lambda)$ in (3.2.5) that uses the QR factorization of J. Equation (3.2.2) is solved by a modification of the two-stage factorization described by Osborne [1972] that allows column pivoting. Subroutine **LMDER** in MINPACK [Moré, Garbow, and Hillstrom (1980)] is an implementation of the method. Variables are scaled internally in **LMDER** according to the following scheme: the initial scaling matrix D_0 is the square root of the diagonal of $J^T J$ evaluated at x_0, and the ith diagonal element of D_k is taken to be the maximum of the ith diagonal element of D_{k-1} and the square root of the ith diagonal element of $J^T J$. Numerical results are presented indicating that this scaling compares favorably with those used

in earlier research. The user also has the option of providing an initial diagonal scaling matrix that is retained throughout the computation.

3.3 Corrected Gauss-Newton Methods

Gill and Murray [1976] propose a linesearch algorithm that divides $\Re^n$ into complementary subspaces $\tilde{\mathcal{R}}$ and $\tilde{\mathcal{N}}$, where $\tilde{\mathcal{R}} \subseteq \mathcal{R}(J^T)$, and $\tilde{\mathcal{N}}$ is nearly orthogonal to $\mathcal{R}(J^T)$. The search direction is the sum of a Gauss-Newton direction in $\tilde{\mathcal{R}}$, and a projected Newton direction in $\tilde{\mathcal{N}}$. This strategy avoids a shortcoming of Gauss-Newton methods — that components of the search direction that are nearly orthogonal to $\mathcal{R}(J^T)$ may not be well determined when J is ill-conditioned — because each component is computed from a reasonably well-conditioned subproblem. The vector $x - x^*$ may become almost entirely in $\mathcal{R}(J^T)$ in a Gauss-Newton method, yet the algorithm computes a search direction that is virtually orthogonal to $\mathcal{R}(J^T)$ due to ill conditioning in the Jacobian (see Fraley [1987b]). Gill and Murray show that both Gauss-Newton algorithms defined by (3.1.4) and Levenberg-Marquardt algorithms generate search directions that lie in $\mathcal{R}(J^T)$, while the Newton search direction generally will have a component in $\mathcal{N}(J)$, the orthogonal complement of $\mathcal{R}(J^T)$, whenever J has linearly dependent columns. For problems with small residuals, they point out that J^TJ is a reasonable approximation to the full Hessian in $\mathcal{R}(J^T)$, but not in $\mathcal{N}(J)$. Thus, in situations where $x - x^*$ is orthogonal to $\mathcal{R}(J^T)$, and J is well-conditioned but has linearly dependent columns (for example, when $m < n$), the Gauss-Newton and Levenberg-Marquardt directions have no component in the direction of $x - x^*$, while Newton's method and also the method of Gill and Murray would have components in both $\mathcal{R}(J^T)$ and $\mathcal{N}(J)$.

A version of this algorithm called the *corrected Gauss-Newton method* [Gill and Murray (1978)] forms the basis for the nonlinear least-squares software in the NAG Library [1984]. Rules based on the relative size of the singular values of J are given for choosing an integer $grade(J)$ to approximate $rank(J)$, and an attempt is made to group together singular values that are similar in magnitude. The method is not as sensitive to $grade(J)$ as Gauss-Newton is to rank estimation, both because of the division of the computation of the search direction into separate components in $\tilde{\mathcal{R}}$ and $\tilde{\mathcal{N}}$, and because $grade(J)$ is varied adaptively based on a measure of the progress of the minimization. Moreover, the rate of convergence is potentially faster than Gauss-Newton or Levenberg-Marquardt methods on problems with nonzero residuals. The quantity $grade(J)$

is reduced when the sum of squares is not adequately decreasing, so that there is the potential of having $\tilde{\mathcal{N}} = \Re^n$ (with exact second derivatives, this implies taking full Newton steps) in the vicinity of a solution.

When $rank(J) = grade(J) = n$, the search direction p is a full-rank Gauss-Newton direction. Otherwise the vector p is computed as the sum of two mutually orthogonal components: a Gauss-Newton direction, and a projected Newton direction. The projected Hessian is replaced by a modified Cholesky factorization (see Section 2.3.1) if it is computationally singular or indefinite. A modified Newton search direction (corresponding to the case $grade(J) = 0$) is used whenever if $|cos(g, p)|$ is smaller than some prescribed value, or if $g^T p$ is positive. A quasi-Newton approximation to B (see the discussion in Section 3.4) and a finite-difference approximation to the projected matrix $Z^T B Z$ along the columns of Z, where Z is an orthogonal basis for $\tilde{\mathcal{N}}$, are given as alternatives to handle cases in which second derivatives of the residual functions are not available or are difficult to compute. Gill and Murray test their method on a set of twenty-three problems, and find that the version that uses quasi-Newton approximations to B does not perform as well as those that use exact second derivatives or finite-difference approximations to a projection of B. They observe only linear convergence for the quasi-Newton version on problems with large residuals. The algorithms are implemented in the NAG Library [1984]; subroutine **E04HEF** uses exact second derivatives, while subroutine **E04GBF** is the quasi-Newton version.

3.4 Special Quasi-Newton Methods

Special quasi-Newton methods for nonlinear least squares use a Hessian of the form $J^T J + \tilde{B}$ in the quadratic model, so that the search direction differs from the Gauss-Newton direction in $\mathcal{R}(J^T)$, and also has a component in $\mathcal{N}(J)$ when J is rank-deficient. The matrix $\tilde{B}$ is a quasi-Newton approximation to the term $B = \sum_{i=1}^{m} \phi_i \nabla^2 \phi_i$ in the Hessian of the nonlinear least-squares objective. Brown and Dennis [1971] first proposed a method in which the Hessian matrix of each of the residuals was updated separately. This approach is impractical because it entails the storage of m symmetric matrices of order n, and more recent research has aimed to approximate B as a sum.

Gill and Murray [1978] discuss a linesearch method in which they use the augmented Gauss-Newton quadratic model only to compute a component of the search direction in a subspace that approximates the null space of the Jacobian (see Section 3.3). They apply the BFGS

formula for unconstrained optimization (see, e. g., Dennis and Moré [1977]) to the matrix $\tilde{H}_k = J_{k+1}^T J_{k+1} + \tilde{B}_k$ with the quasi-Newton condition

$$\tilde{H}_{k+1} s_k = y_k$$

where

$$s_k \equiv x_{k+1} - x_k \quad \text{and} \quad y_k \equiv g_{k+1} - g_k,$$

and then form $\tilde{B}_{k+1} = \tilde{H}_{k+1} - J_{k+1}^T J_{k+1}$. They point out that, if $J_{k+1}^T J_{k+1} + \tilde{B}_k$ is positive definite, and $y_k^T s_k > 0$, then $J_{k+1}^T J_{k+1} + \tilde{B}_{k+1}$ is also positive definite with this scheme. In order to safeguard the method, the projected approximate Hessian is replaced by a modified Cholesky factorization when it is singular or indefinite. In addition, if the cosine of the angle between the search direction and the gradient of the nonlinear least-squares objective exceeds a fixed threshold value, a modified Newton step with the full augmented approximate Hessian is taken. See Section 3.3 for a summary of their observations on the performance of the methods.

Dennis, Gay, and Welsch [1981a] apply a scaled DFP update to $\tilde{B}_k$ at each step. The new approximation $\tilde{B}_{k+1}$ solves

$$\min_{B;H} \|H^{-1/2}(\tau_k \tilde{B}_k - B)H^{-1/2}\|_F$$

subject to

$$H s_k = y_k; \quad H \text{ positive definite}$$

$$B s_k = J_{k+1}^T f_{k+1} - J_k^T f_{k+1}; \quad B \text{ symmetric},$$

where

$$\tau_k \equiv \min\{|y_k^T s_k / s_k^T \tilde{B}_k s_k|, 1\}.$$

The scale factor τ_k is based on the observation that the quasi-Newton approximation to B is often too large with the unscaled update, on account of the contribution of the residuals. The term $|y_k^T s_k / s_k^T \tilde{B}_k s_k|$ in τ_k is derived from the self-scaling principles for quasi-Newton methods of Oren [1973], and attempts to shift the eigenvalues of the approximation $\tilde{B}_k$ to overlap with those of B_k, using curvature information at x_k. The algorithm forms the basis for the ACM computer program **NL2SOL** [Dennis, Gay, and Welsch (1981b)], which is distributed by the PORT Library [1984] as subroutines **N2G** and **DN2G**. It is implemented as an adaptive method, in that Gauss-Newton steps are taken if the Gauss-Newton quadratic model predicts the reduction in the function better than the quadratic model that includes the term involving $\tilde{B}$. A trust-region

strategy is used to *enforce* global convergence. Numerical results are given in Dennis, Gay, and Welsch [1981a] for a set of twenty-four test problems, many with two or three different starting values.

4. Discussion and Summary of Numerical Results

In this section, we summarize numerical results obtained for the unconstrained optimization and nonlinear least-squares methods; *more detailed results are tabulated in the Appendix.* The tests were performed using the following software:

- **DMNG/SUMSOL** - Trust-region method for unconstrained optimization that uses a quasi-Newton approximation to the Hessian matrix. From the PORT Library [1984].
- **NPSOL** - Linesearch method for unconstrained optimization that uses a quasi-Newton approximation to the Hessian matrix. From the Systems Optimization Laboratory, Stanford University (see Gill et al. [1986b; 1987]). Also available in the NAG Library.
- **DMNH/HUMSOL** - Trust-region method for unconstrained optimization that uses analytic second derivatives. From the PORT Library [1984].
- **MNA** - Linesearch method for unconstrained optimization that uses analytic second derivatives. This implementation, which is available at Stanford Linear Accelerator Center, is from the National Physical Laboratory, England. It is essentially the same as subroutine **E04LBF** from the NAG Library [1984].
- G-N - Gauss-Newton algorithm for nonlinear least squares that uses **LSSOL** (Gill et al. [1986a]) to solve the linear least squares subproblems, and the linesearch from **NPSOL** (Gill et al. [1986b]). Both **LSSOL** and **NPSOL** are also available in the NAG Library.
- **LMDER** - Levenberg-Marquardt method for nonlinear least squares. From MINPACK (Moré, Garbow, and Hillstrom [1980]).
- **DN2G/NL2SOL** - Adaptive method for nonlinear least squares (combines Gauss-Newton, Levenberg-Marquardt, and special quasi-Newton methods). From the PORT Library [1984].
- **LSQFDQ** - Corrected Gauss-Newton method that uses a quasi-Newton approximation to the Hessian matrix. This implementation, which is available at Stanford Linear Accelerator Center, is from the National Physical Laboratory, England. It is essentially the same as subroutine **E04GBF** from the NAG Library [1984].
- **LSQSDN** - Corrected Gauss-Newton method that uses analytic second derivatives. This implementation, which is available at Stanford Linear Accelerator Center, is from the National Physical Laboratory, England. It is essentially the same as subroutine **E04HEF** from the NAG Library [1984].

Information about the individual test problems is given in the Appendix. The number of function evaluations required by each subroutine is listed in the tables at the end of this section. In addition, the following symbols are used:

^{0}	-	zero-residual problem
^{L}	-	linear least-squares problem
—	-	failure to achieve an approximate solution
~	-	appears to be unable to terminate at an approximate solution
^{l}	-	local minimum
^{t}	-	termination criteria satisfied at a point away from a local minimum
^{s}	-	failed with default step length or trust-region size

Two columns of figures corresponding to two different values of a single parameter are given for each subroutine. For the Gauss-Newton methods, the parameter affects rank estimation; for all of the other methods, the parameter affects termination criteria. See the tables of numerical results given in the Appendix for information about the precise choices that were made. The wide variability in the numerical results makes it difficult to draw definitive conclusions about the relative performance of the software, because observations of small samples could result in misleading generalizations. The sources of this variability are discussed below.

First, the number of function evaluations may not be an adequate basis for comparison. The routines vary in the number of gradient evaluations performed per function evaluation, and second-derivative methods require evaluation of the Hessian matrix. Moreover, when function evaluations are relatively inexpensive, costs could be dominated by other portions of the computation. Another difficulty in making comparisons is that the definition of an acceptable minimum varies from routine to routine. For example, the norm of the gradient of the nonlinear least-squares objective, $||g||$, at an alleged solution x^* may differ considerably for different software, although $g(x^*) = 0$ is a necessary condition for a minimum. (On problem 10., **LMDER** terminates at a point for which $||g||$ is of order 10^0, while **DN2G** terminates at a point for which $||g||$ is of order 10^{-3}.) Most algorithms do not attempt to reduce $||g||$ directly, but convergence criteria may include a threshold on $||g(x^*)||$. Depending on how this threshold is used in relation to other criteria, some routines may spend more function evaluations in anticipation of a reduction in $||g||$ than others. A small value of $||g||$ means greater certainty that a minimum has actually been obtained, but may be unreasonably expensive to achieve in practice.

Second, aside from design choices that define a particular implementation of an algorithm, the user is permitted to specify certain parameters that may affect *performance*. Fraley [1987b] gives examples that illustrate the sensitivity of Gauss-Newton methods to rank estimation criteria (see, e. g., problems **35b.**, **36a.**, and **20d.**). For problems on which an algorithm is linearly convergent, small changes in tolerances that are used to define convergence criteria can mean substantial differences in the amount of computation required in order to obtain a point satisfying conditions for convergence (see, e. g., **DNMG** on **24b.**, and **LMDER** on **40.**). Selection of a maximum steplength or an initial trust-region radius can also be critical factor in the performance of a method. In these tests, the default values for these parameters were altered only in cases where a method was initially observed to fail by attempting to evaluate problem functions outside the region in which they are numerically defined (see, e. g., the results for the DeVilliers and Glasser test problems **42.** and **43.**). Failures of this sort may be caused either by poor scaling among the variables in the problem, or by ill-conditioning within subproblems. Hiebert [1981] is of the opinion that failure due to ill-conditioning can be avoided in software, but that it is not always possible to anticipate abnormal terminations that are caused by bad scaling.

In **NPSOL**, one can specify bounds on the variables, as well as adjust the maximum step length, in order to deal with this type of difficulty. Subroutine **E04LBF**, the version of **NMA** that is available in the NAG Library [1984], also provides for bounds on variables, and there are alternative versions of all of the PORT software used in these tests that allow bounds to be specified. Unfortunately, when bounds on the variables are included in the problem formulation, local minima at which the bounds are active may be found rather than local minima for the nonlinear least-squares problem. See Hanson and Krogh [1987] for numerical results in which simple bounds are included for some problems. Holt and Fletcher [1979] give an algorithm designed for nonlinear least-squares problems with explicit bounds on the variables.

Third, the performance of any given method over the set of test problems is by no means uniform, and it is not easy to separate the problems into classes for which the behavior of an algorithm can be categorized. One reason for this is that many of the test problems recur in the literature precisely because they have certain distinguishing properties. Powell's singular function and variants (**13.** and **22.**) are zero-residual problems in which the Jacobian becomes singular at the solution. The McKeown test problems (**39.**, **40.**, and **41.**) are chosen so that the Jacobian is well-conditioned everywhere, and the rate of convergence for the Gauss-Newton method with unit steplength can be controlled by varying a single parameter (the parameter can

also be chosen so that the *unit-step* Gauss-Newton method *diverges*). Both Powell's singular function and McKeown's test problems are constructed analytically rather than derived from data-fitting applications. The matrix square root problems (36.) are examples of small, dense, nonlinear systems of equations requiring a very accurate solution. Watson's function (20.) comes from polynomial interpolation, and has multiple local minima with small, but nonzero, residuals. It also has the feature that the Jacobian becomes increasingly ill-conditioned as the problem size is increased (see Fraley [1987b]). The Gulf Research and Development function (11.) has discontinuities in the derivative of each residual on a one-dimensional subspace and hence violates the assumption (made in developing all of the algorithms we have discussed) that the sum of squares has continuous second derivatives. The results for the DeVilliers and Glasser test problems (42. and 43.) illustrate variability in performance due to the use of different starting values. More generally, the behavior of a given method for a certain type of residual function may not be uniform over several sets of defining data of similar magnitude, as shown by the results for the Dennis, Gay, and Vu test problems (44. and 45.).

Finally, there is considerable diversity in performance among the routines tested, and few generalizations are possible. Our data generally supports the use of nonlinear least-squares software over that designed for general unconstrained minimization, but there are some exceptions (see, e. g., the McKeown test problems 39. - 41.). Of the nonlinear least-squares routines, DN2G (NL2SOL) is often the best (the Dennis, Gay, and Vu test problems 44. and 45. are examples of exceptions). When second derivatives are relatively cheap to obtain, the use of an unconstrained optimization method that uses exact second derivatives may be a reasonable alternative to a nonlinear least-squares method (see, e. g., the results for the penalty function 23b.). Our tests do not indicate overall superiority of any particular method over the others; in situations in which a variety of problems are being solved, we conclude that it is desirable to have the flexibilty to choose from among several methods.

Summary of Results : Unconstrained Optimization Methods

(number of function evaluations)

Moré, Garbow, and Hillstrom Test Problems

	n	m	DMNG		NPSOL		DMNH		MNA		
1.^{0}	2	2	40	42	27	29	32	32	14	14	
2.^{0}	2	2	12^l	12^l	9	10	10^l	10^l	8	8	
3.^{0}	2	2	217	220	897	897	130	132	175	175	
4.^{0}	2	3	66	67	20	20	22	23	1	1	
5.^{0}	2	3	16	17	20	22	11	12	35^s	35^s	
6.	2	10	33	34	14^s	15^s	11	11	12	12	
7.^{0}	3	3	28	30	37	38	16	17	14	14	
8.	3	15	19	22	22	23	9	10	11	11	
9.	3	15	8	12	8	9	4	4	3	3	
10.	3	16	465	467	450	450	382	388	249	249	
11.^{0}	3	10	4^l	327	2^l	2^l	290	292	538	538	
12.^{0}	3	10	43	45	34	35	24	24	43	43	
13.^{0}	4	4	62^t	89	66	71	27^t	38	38	23^t	23^t
14.^{0}	4	6	100	102	50	51	42	49	54	54	
15.	4	11	35	36	33	35	11	12	20	20	
16.	4	20	46	47	24	25	11	13	9	10	
17.	5	33	69	72	32^{sl}	56^s	46^s	47^s	43	43	
18.^{0}	6	13	45	47	45^l	45^l	–	–	44	44	
19.	11	65	69	72	88	90	23	24	7^s	8^s	
20a.	6	31	35	37	43	46	15	15	13	13	
20b.	9	31	76	79	83	85	20	22	14	14	
20c.	12	31	89^l	148^l	55^l	151	24	24	14	14	
20d.	20	31	110^l	134^l	73^l	114^l	50^l	~	1295^l	1295^l	
21a.^{0}	10	10	120	125	101	104	25	26	14	14	
21b.^{0}	20	20	189	193	252	265	27	27	14	14	
22a.^{0}	12	12	143^t	235	83^t	165^t	28^t	40	23^t	23^t	
22b.^{0}	20	20	187^t	344	103^t	196^t	29^t	40	24^t	24^t	
23a.	4	5	77	78	198	198	42	43	43	43	
23b.	10	11	80	81	117	124	44	45	44	44	
24a.	4	8	364	472	23^l	462	126	128	158	158	
24b.	10	20	475	632	368	419	158	162	133	133	

Summary of Results : Nonlinear Least-Squares Methods

(number of function evaluations)

Moré, Garbow, and Hillstrom Test Problems

	G-N		LMDER		DN2G		LSQFDQ		LSQSDN	
$1.^{0}$	31	31	22	22	14	14	31	31	31	31
$2.^{0}$	180^{l}	235^{l}	14^{l}	21^{l}	10^{l}	12^{l}	36^{l}	36^{l}	18^{l}	18^{l}
$3.^{0}$	31^{l}	42	19	19	64	65	47	47	47	47
$4.^{0}$	54	54	40^{t}	54	40^{t}	53	64	64	53	53
$5.^{0}$	8	8	9	10	9	10	14	14	10	10
6.	$-^{s}$	$-^{st}$	21	28	14	16	54	54	36^{s}	36^{s}
$7.^{0}$	13	13	11	12	13	14	20	20	14	14
8.	7	7	6	7	7	8	13	13	6	6
9.	3	3	4	5	3	5	3	3	3	3
10.	$-^{t}$	30	126	126	132	133	18	18	17	17
$11.^{0}$	–	–	–	–	–	–	69^{l}	69^{l}	30^{sl}	30^{sl}
$12.^{0}$	7	7	7	8	8	9	12	12	8	12
$13.^{0}$	16^{t}	16^{t}	65	65	19^{t}	25	18^{t}	18^{t}	18^{t}	18^{t}
$14.^{0}$	96	96	70	70	52	53	81	81	87	99
15.	43	43	18	28	11	12	30	30	16	16
16.	3651	3651	264	356	21	22	62	62	45	45
17.	13	13	18	19	26	27	19	19	14	18
$18.^{0}$	$-^{s}$	–	46	46	45	46	$-^{s}$	$-^{s}$	243^{s}	247^{s}
19.	24	24	17	19	20	22	33	33	19	19
20a.	12	12	8	10	13	13	32	32	9	9
20b.	6	6	9	10	12	15	6	6	6	6
20c.	6	6	10	12	14	14	6	6	6	6
20d.	6^{l}	6	18	23	7^{l}	~	18	18	10	12
$21a.^{0}$	31	31	22	22	27	27	31	31	31	31
$21b.^{0}$	31	31	22	22	16	16	31	31	31	31
$22a.^{0}$	16^{t}	16^{t}	72	72	20^{t}	26	18^{t}	18^{t}	18^{t}	18^{t}
$22b.^{0}$	16^{t}	16^{t}	69	69	19^{t}	26	18^{t}	18^{t}	18^{t}	18^{t}
23a.	90	90	34	44	36	37	80	80	58	58
23b.	274	274	84	104	61	68	143	143	124	124
24a.	1043	1043	151	156	139	142	411	411	228	228
24b.	–	–	80	88	129	138	566^{s}	566^{s}	150	150

Summary of Results : Unconstrained Optimisation Methods

(number of function evaluations)

Moré, Garbow, and Hillstrom Test Problems (continued)

	n	m	DMNG		NPSOL		DMNH		MNA	
25a.^{0}	10	12	20	21	19	20	15	15	14	14
25b.^{0}	20	22	25	26	24	24	18	19	17	18
26a.^{0}	10	10	34^{l}	37^{l}	33^{l}	35^{l}	11^{l}	12^{l}	21	21
26b.^{0}	20	20	62^{l}	65^{l}	76^{l}	78^{l}	20^{l}	20^{l}	30	30
27a.^{0}	10	10	13	16	18	19	9	10	22	22
27b.^{0}	20	20	15	18	30	32	11	12	31	31
28a.^{0}	10	10	31	34	33	36	4	4	4	4
28b.^{0}	20	20	60	64	54	56	4	4	4	4
29a.^{0}	10	10	8	10	7	8	4	5	4	4
29b.^{0}	20	20	8	10	7	8	4	5	4	4
30a.^{0}	10	10	51	57	37	40	6	7	7	7
30b.^{0}	20	20	65	88	65^{l}	67^{l}	6	7	7	7
31a.^{0}	10	10	46	60	67^{l}	70^{l}	9	9	9	9
31b.^{0}	20	20	47	63	141	144	9	9	9	9
32.^{L}	10	20	6	6	2	2	6	6	4	4
33.^{L}	10	20	4	4	4	4	5	5	27	27
34.^{L}	10	20	5	5	4	4	6	6	20	20
35a.	8	8	34	38	31	33	14	14	41	41
35b.^{0}	9	9	44	46	29	32	17	18	66	66
35c.	10	10	41^{l}	45^{l}	37^{l}	42^{l}	19	20	86	86

Matrix Square Root Test Problems

	n	m	DMNG		NPSOL		DMNH		MNA	
36a.^{0}	4	4	–	–	–^{t}	–^{t}	–	–	–	–
36b.^{0}	9	9	–	–	–^{t}	–^{t}	–	–	–	–
36c.^{0}	9	9	69	101	3	3	31	35	3188	3188
36d.^{0}	9	9	–	–	–^{t}	–^{t}	–	–	–	–

Summary of Results : Nonlinear Least-Squares Methods

(number of function evaluations)

Moré, Garbow, and Hillstrom Test Problems (continued)

	G-N		LMDER		DN2G		LSQFDQ		LSQSDN	
25a.^{0}	11	11	11	12	15	16	16	16	12	17
25b.^{0}	13	13	13	14	19	19	18	18	14	19
26a.^{0}	16	16	28^{l}	37^{l}	11	12	22	22	18	22
26b.^{0}	20	20	57^{l}	71^{l}	39	42	18	24	18	26
27a.^{0}	21	21	15	15	8	9	26^{s}	26^{s}	22^{s}	28^{s}
27b.^{0}	31^{s}	31^{s}	5^{l}	18	11	12	31^{s}	31^{s}	21^{s}	27^{s}
28a.^{0}	4	4	5	5	4	4	4	4	4	4
28b.^{0}	4	4	5	5	3	4	4	4	4	4
29a.^{0}	4	4	5	5	4	4	10	10	6	6
29b.^{0}	4	4	5	5	4	4	10	10	6	6
30a.^{0}	6	6	6	7	8	9	11	11	7	11
30b.^{0}	6	6	6	7	8	9	11	11	7	11
31a.^{0}	7	7	7	8	10	11	12	12	8	12
31b.^{0}	7	7	7	8	10	11	12	12	8	12
32.^{L}	2	2	3	3	5	5	2	2	2	2
33.^{L}	3	–	3	8	18	18	2	2	2	2
34.^{L}	3	3	3	7	13	13	2	2	2	2
35a.	–^{t}	–	40	53	23	24	87^{s}	87^{s}	74	74
35b.^{0}	148	–^{t}	12	13	11	11	34	34	30	34
35c.	–	–^{t}	25	34	17^{l}	19^{l}	73^{sl}	73^{sl}	43^{l}	43^{l}

Matrix Square Root Test Problems

	G-N		LMDER		DN2G		LSQFDQ		LSQSDN	
36a.^{0}	2885^{l}	36	–	–	–	–	44	44	38	38
36b.^{0}	683^{l}	36	9	10	16^{l}	–	44	44	38	38
36c.^{0}	20	20	29	40	16	22	28	28	28	28
36d.^{0}	74^{l}	–	2	2	4	4	424	424	380	380

Summary of Results : Unconstrained Optimization Methods

(number of function evaluations)

Hanson Test Problems

	n	*m*	DMNG		NPSOL		DMNH		MNA	
37.	2	16	22	22	14	15	16	17	6	6
38.	3	16	31	32	21*	23*	14	14	13	13

McKeown Test Problems

	n	*m*	DMNG		NPSOL		DMNH		MNA	
39a.	2	3	9	11	10	11	4	4	4	4
39b.	2	3	9	10	9	10	4	4	4	4
39c.	2	3	6	7	6	8	4	5	4	4
39d.	2	3	8	9	10	11	6	6	6	6
39e.	2	3	11	12	16	17	8	8	8	8
39f.	2	3	11	11	28	29	11	11	11	11
39g.	2	3	17	18	30	31	13	14	14	14
40a.	3	4	11	12	11	12	4	4	4	4
40b.	3	4	10	12	11	12	4	5	4	4
40c.	3	4	9	10	9	10	5	5	5	5
40d.	3	4	10	11	14	15	5	6	6	6
40e.	3	4	11	13	19	20	7	7	8	8
40f.	3	4	14	16	33	34	10	10	10	10
40g.	3	4	18	20	45	46	13	13	13	13
41a.	5	10	11	13	12	12	4	4	4	4
41b.	5	10	11	13	12	13	4	4	4	4
41c.	5	10	13	14	12	14	8	8	8	8
41d.	5	10	17	20	17	20	8	9	9	9
41e.	5	10	24	26	51	54	11	12	12	12
41f.	5	10	27	31	51	53	14	14	14	14
41g.	5	10	32	35	62	69	17	17	17	17

Summary of Results : Nonlinear Least-Squares Methods

(number of function evaluations)

Hanson Test Problems

	G-N		LMDER		DN2G		LSQFDQ		LSQSDN	
37.	39	39	15	21	10	11	25	25	9	9
38.	58	58	18	28	10	12	30	30	10	10

McKeown Test Problems

	G-N		LMDER		DN2G		LSQFDQ		LSQSDN	
39a.	8	8	5	6	5	5	17	17	4	4
39b.	32	32	14	21	6	7	24	24	6	6
39c.	23	23	18	25	7	8	22	23	9	9
39d.	681	681	20	28	7	8	31	32	12	12
39e.	–	–	28	44	9	10	32	32	12	12
39f.	–	–	31	44	14	15	43	43	25	25
39g.	–	–	39	44	18	20	49	49	39	39
40a.	13	13	6	9	7	7	18	18	5	5
40b.	16	16	14	17	7	11	19	19	6	6
40c.	380	380	16	22	9	10	27	27	11	11
40d.	781	781	26	40	9	9	33	34	13	13
40e.	–	–	90	146	10	11	70	72	45	45
40f.	–	–	180	272	13	14	92	92	49	49
40g.	–	–	206	319	23	25	123*	123*	85	85
41a.	5	5	4	4	4	4	8	8	4	4
41b.	6	6	4	5	4	5	18	18	4	4
41c.	12	12	6	8	6	6	21	21	5	5
41d.	30	30	15	22	9	11	38	38	9	9
41e.	222	222	29	38	17	20	47	47	14	14
41f.	933	933	57	89	24	27	54	54	16	16
41g.	3285	3285	84	144	29	30	62	62	21	21

Summary of Results : Unconstrained Optimization Methods .

(number of function evaluations)

DeVilliers and Glasser Test Problems

	n	m	DMNG		NPSOL		DMNH		MNA	
42a.^{0}	4	24	53	56	60^{s}	61^{s}	28	28	41^{s}	41^{s}
42b.^{0}	4	24	103^{s}	104^{s}	140^{s}	141^{s}	35	36	16	16
42c.^{0}	4	24	76	78	51^{s}	52^{s}	30	31	6	6
42d.^{0}	4	24	61	64	56^{s}	57^{s}	30	30	6	6
43a.^{0}	5	16	49	51	44^{sl}	53^{sl}	22	22	30^{s}	30^{s}
43b.^{0}	5	16	58	60	37^{sl}	38^{sl}	26	27	17^{s}	17^{s}
43c.^{0}	5	16	41	44	44^{sl}	54^{sl}	21	21	89^{s}	89^{s}
43d.^{0}	5	16	57	60	112^{sl}	120^{sl}	27^{s}	28^{s}	41^{s}	42^{s}
43e.^{0}	5	16	51	53	95^{s}	97^{s}	28^{s}	29^{s}	142^{s}	143^{s}
43f.^{0}	5	16	45	48	56^{s}	59^{s}	17	18	37^{s}	37^{s}

Dennis, Gay, and Vu Test Problems

	n	m	DMNG		NPSOL		DMNH		MNA	
44a.^{0}	6	6	441	444	488	490	179	180	143	144
44b.^{0}	6	6	31	34	57	59	9	10	46	46
44c.^{0}	6	6	3726	3731	–	–	194	195	914	915
44d.^{0}	6	6	–^{t}	3865	–	–	187	188	915	916
44e.^{0}	6	6	–^{t}	2815	1976	1978	219	220	475	476
45a.^{0}	8	8	284	288	474	476	63	64	186	186
45b.^{0}	8	8	36	40	82	84	15	16	38	38
45c.^{0}	8	8	6197	6200	–	–	321	322	1416	1416
45d.^{0}	8	8	7929	7934	1654	1656	328	329	1478	1479
45e.^{0}	8	8	3341	3346	–	–	351	352	1441	1441

Summary of Results : Nonlinear Least-Squares Methods

(number of function evaluations)

DeVilliers and Glasser Test Problems

	G-N		LMDER		DN2G		LSQFDQ		LSQSDN	
42a.^{0}	51^s	51^s	18	19	29^s	29^s	73^{sl}	73^{sl}	60^{sl}	64^{sl}
42b.^{0}	611^s	–	48^s	49^s	74^s	74^s	94^s	94^s	75^s	75^s
42c.^{0}	27^s	27^s	20^s	20^s	32^s	32^s	46^s	46^s	48^s	48^s
42d.^{0}	24^s	24^s	15^s	16^s	23	24	27^s	27^s	27^s	27^s
43a.^{0}	23^s	23^s	14^s	15^s	31	32	33^s	33^s	25^s	33^s
43b.^{0}	15^{sl}	20^s	18^s	18^s	20	20	45^s	45^s	37^s	45^s
43c.^{0}	24^s	24^s	11^s	11^s	34^l	41^l	33^s	33^s	25^s	33^s
43d.^{0}	18^s	18^s	22^s	23^s	17	17	38^s	38^s	30^s	38^s
43e.^{0}	31^s	31^s	12^s	13^s	28	29	27^s	27^s	19^s	27^s
43f.^{0}	22	22	12	13	20	20	31	31	23	31

Dennis, Gay, and Vu Test Problems

	G-N		LMDER		DN2G		LSQFDQ		LSQSDN	
44a.^{0}	171	171	37	38	58	59	97	97	93	95
44b.^{0}	5	5	5	6	7	7	10	10	6	10
44c.^{0}	55	55	108	109	93	94	47	47	47	47
44d.^{0}	35	35	98	99	97	98	40	40	40	40
44e.^{0}	42	42	82	83	83	83	47	47	47	47
45a.^{0}	171	171	47	48	65	66	97	97	93	95
45b.^{0}	5	5	5	6	8	8	12	12	6	12
45c.^{0}	41	41	164	165	129	130	47	47	47	47
45d.^{0}	35	35	144	145	168	168	42	42	42	42
45e.^{0}	42	42	130	131	173	173	49	49	43	43

5. Appendix : Numerical Results

In this section, numerical results are presented for a large set of test problems, using software based on the unconstrained optimization techniques and methods for nonlinear least squares problems discussed in Sections 2 and 3.

5.1 Sources and Presentation

The following is a list of the software sources that were used to obtain the results:

subroutine	source	problem type	derivatives
DMNG/SUMSOL	PORT	unconstrained	first
NPSOL	SOL / NAG	unconstrained	first
DMNH/HUMSOL	PORT	unconstrained	second
MNA	NPL / NAG	unconstrained	second
G-N	uses SOL / NAG LSSOL	least squares	first
LMDER	MINPACK	least squares	first
DN2G/NL2SOL	PORT	least squares	first
LSQFDQ	NPL / NAG	least squares	first
LSQSDN	NPL / NAG	least squares	second

ACM	-	Association for Computing Machinery
MINPACK	-	Argonne National Laboratory, U. S. A.
NAG	-	Numerical Algorithms Group
NPL	-	National Physical Laboratory, England
PORT	-	PORT Mathematical Software Library, A. T. & T. Bell Laboratories, Inc.
SOL	-	Systems Optimization Laboratory, Stanford University

All of the programs were run in FORTRAN using double precision on the IBM 3081 and IBM 3033 computers at Stanford Linear Accelerator Center, for which

relative machine precision $\epsilon_M = 2.22\ldots \times 10^{-16}$; $\sqrt{\epsilon_M} = 1.49\ldots \times 10^{-8}$.

In the tables, associated with the label 'est. err.', we include the quantity

$$\frac{\|f^*\|_2^2 - \|f_{best}\|_2^2}{1 + \|f_{best}\|_2^2}, \tag{5.1.1}$$

where f^* is the value of f at the point of termination, and $\|f_{best}\|_2$ is the best available estimate of the norm of the solution, in order to get some idea of the error in $\|f^*\|_2$. For those problems

that have nonzero residuals, the value of $||f_{best}||_2$ is given to six figures of accuracy, rounded down.

The following abbreviations are used in the headings of the tables:

est. err.	-	error estimate (5.1.1)
conv.	-	termination conditions

The following abbreviations are used in the tables to describe conditions under which the algorithm terminates abnormally:

F LIM.	-	function evaluation limit reached
TIME	-	time limit exceeded
LOOP	-	subroutine appears to loop

For information on the test problems, see Section 5.9.

5.2 Trust-Region Methods

(PORT/ACM DMNH/HUMSL and DMNG/SUMSL)

5.2.1 Software and Algorithms

The results were obtained using subroutines DMNH and DMNG, which are double precision versions of the ACM algorithms HUMSL and SUMSL available in the PORT Library [1984]. A subproblem of the form

$$\min_{p \in \Re^n} Q_k(p) \equiv g_k^T p + \frac{1}{2} p^T H_k p$$

$$\text{subject to} \quad \|D_k p\|_2 \leq \delta_k$$

is solved at each iteration for the step p_k to the next iterate, where D_k is a diagonal scaling matrix, and H_k is the exact Hessian matrix at x_k in DMNH, and a quasi-Newton approximation in DMNG (see Sections 2.3.2 and 2.4).

5.2.2 Parameters

Parameters were kept at their default values with the following exceptions:

IV(MXFCAL)	-	min {9999, 1000n}	function evaluation limit
IV(MXITER)	-	min {9999, 1000n}	iteration limit
V(AFCTOL)	-	TOL^2 (varied; see tables)	absolute function convergence tolerance
V(RFCTOL)	-	TOL (varied; see tables)	relative function convergence tolerance
V(SCTOL)	-	ϵ_M	singular convergence tolerance
V(XCTOL)	-	TOL (varied; see tables)	x convergence tolerance
V(XFTOL)	-	ϵ_M	false convergence tolerance
V(LMAXO)	-	usually 1.0 (default) †	initial trust-region diameter

† In some cases the default V(LMAXO) = 1.0 for the initial diameter of the trust-region was too large and overflow occurred during function evaluation. These cases are indicated in the table by giving the lower value of V(LMAXO) that was subsequently used to obtain the results in the column labeled "init. diam.".

See Dennis, Gay, and Welsch [1981a, 1981b], Gay [1983], and PORT [1984] for details concerning the parameters.

5.2.3 Convergence Criteria

The following quantities will be used in describing the convergence criteria :

objective function	:	$\mathcal{F}_k \ \ (= \frac{1}{2} f_k^T f_k)$
objective gradient	:	$g_k = \nabla \mathcal{F}_k \ \ (= J_k^T f_k)$
current step	:	p_k, the minimizer of the subproblem
Newton step	:	$p_N \begin{cases} H_k^{-1} g & \text{if } H_k \text{ is positive definite;} \\ \textit{undefined} & \text{otherwise.} \end{cases}$
Newton reduction	:	$\rho_N = \begin{cases} -Q_k(p_N) & \text{if } H_k \text{ is positive definite;} \\ 0 & \text{otherwise.} \end{cases}$
predicted reduction	:	$\rho_r = -Q_k(p_k)$
actual reduction	:	$\rho_A = \mathcal{F}_k - \mathcal{F}(x_k + p_k)$
scaled distance	:	$\nu(x, y, D) = \dfrac{\max_{1 \le i \le n} \{\|(D(x-y))_i\|\}}{\max_{1 \le i \le n} \{\|(Dx)_i\| + \|(Dy)_i\|\}}$. †

† Here v_i denotes the ith component of the vector v. There is a provision for the user to replace the function ν; we used the default in all of the tests.

The convergence criteria used in DMNH and DMNG are as follows:

- *Absolute function convergence* occurs at x_k if

$$|\mathcal{F}_k| < \texttt{V(AFCTOL)}. \tag{5.2.1}$$

- *Relative function convergence* is intended to approximate the condition

$$\mathcal{F}_k - \mathcal{F}(x^*) \le \texttt{V(RFCTOL)}\,|\mathcal{F}_k|.$$

The test actually used is

$$\rho_N \le \texttt{V(RFCTOL)}\,|\mathcal{F}_k|. \tag{5.2.2}$$

- *x convergence* is intended to approximate the condition

$$\nu(x_k, x^*, D_k) \le \texttt{V(XCTOL)},$$

The test actually used is

$$p_k = p_N \ \text{ and } \ \nu(x_k, x_k + p_k, D_k) \le \texttt{V(XCTOL)}. \tag{5.2.3}$$

- *Singular convergence* is intended to approximate the condition

$$\mathcal{F}_k - \min\,\{\mathcal{F}(y) \mid \|D_k(y - x_k)\| \le \texttt{V(LMAXS)}\} < \texttt{V(SCTOL)}\,|\mathcal{F}_k|,$$

where D_k is the diagonal scaling matrix at the kth iterate. The test for singular convergence is made only when none of the convergence criteria listed above holds. It is meant to indicate

relative function convegence when the Hessian in the subproblem is singular.

The actual test is

$$\mathcal{F}_k - \min \{Q_k(y) \mid \|D_k(y - x_k)\| \le \mathtt{V(LMAXS)}\} < \mathtt{V(SCTOL)}\,|\mathcal{F}_k|. \qquad (5.2.4)$$

Under certain conditions, the test (5.2.4) is repeated for a step of length V(LMAXS).

• *False convergence* is returned if none of the other criteria are satisfied and a trial step no larger than V(XFCTOL) is rejected. This usually indicates either an error in computing the objective gradient, a discontinuity (in $\mathcal{F}$ or g) near the current iterate, or that one or more of the convergence tolerances (V(RFCTOL), V(XCTOL), and V(AFCTOL)) are too small relative to the accuracy to which the objective is computed.

The test actually used is

$$\mathcal{F}_k - \mathcal{F}(x_k + p_k) \le \mathtt{V(TUNER1)}\,\rho_r \quad \text{and} \quad \nu(x_k, x_k + p_k, D_k) \le \mathtt{V(XFTOL)}. \qquad (5.2.5)$$

where the parameter V(TUNER1) is adjustable, although in these tests the default value 0.1 is used throughout.

Except for (5.2.1), tests for convergence are performed only when

$$\rho_A \le 2\rho_r. \qquad (5.2.6)$$

See Dennis, Gay, and Welsch [1981a, 1981b], Gay [1983], and PORT [1984] for more discussion of the convergence criteria.

The following abbreviations are used in the tables to describe the conditions under which the algorithm terminates:

ABS. F	-	(5.2.1)
REL. F	-	(5.2.2) and (5.2.6)
X	-	(5.2.3) and (5.2.6)
X, F	-	(5.2.2) and (5.2.3) and (5.2.6)
SING.	-	(5.2.4) and (5.2.6)
FALSE	-	(5.2.5) and (5.2.6)
F LIM.	-	function evaluation limit reached
TIME	-	time limit exceeded
LOOP	-	subroutine appears to loop

The total number of Jacobian evaluations is either equal to the total number of iterations of the method, or it is one more than the number of iterations. The number in the column labeled "iters. / J evals." is followed by a "+" if an extra Jacobian evaluation was used in the computation.

Numerical Results for DMNG

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
1.[0]	2	2	10^{-8}		40	35+	1.41	10^{-9}	10^{-8}	10^{-19}	ABS. F
			10^{-12}		42	37+	1.41	10^{-14}	10^{-14}	10^{-29}	X
2.[0]	2	2	10^{-8}		12	10+	11.4	10^{1}	10^{-7}	10^{1}	REL. F
			10^{-12}		12	10+	11.4	10^{1}	10^{-7}	10^{1}	REL. F
3.[0]	2	2	10^{-8}		217	160+	9.11	10^{-9}	10^{-4}	10^{-16}	ABS. F
			10^{-12}		220	162+	9.11	10^{-12}	10^{-7}	10^{-25}	ABS. F
4.[0]	2	3	10^{-8}		66	11+	10^{6}	10^{-8}	10^{-2}	10^{-15}	X
			10^{-12}		67	12+	10^{6}	10^{-11}	10^{-8}	10^{-22}	X
5.[0]	2	3	10^{-8}		16	14+	3.04	10^{-11}	10^{-10}	10^{-22}	ABS. F
			10^{-12}		17	15+	3.04	10^{-14}	10^{-13}	10^{-28}	ABS. F
6.	2	10	10^{-8}		33	20+	.365	10^{-1}	10^{-4}	10^{-6}	REL. F
			10^{-12}		34	21+	.365	10^{-1}	10^{-7}	10^{-6}	REL. F
7.[0]	3	3	10^{-8}		28	23+	1.00	10^{-10}	10^{-8}	10^{-19}	ABS. F
			10^{-12}		30	25+	1.00	10^{-14}	10^{-13}	10^{-28}	ABS. F
8.	3	15	10^{-8}		19	17+	2.60	10^{-1}	10^{-7}	10^{-8}	REL. F
			10^{-12}		22	20	2.60	10^{-1}	10^{-12}	10^{-8}	REL. F
9.	3	15	10^{-8}		8	5+	1.08	10^{-4}	10^{-11}	10^{-14}	X, REL. F
			10^{-12}		12	9	1.08	10^{-4}	10^{-17}	10^{-14}	REL. F
10.	3	16	10^{-8}		465	325+	10^{4}	10^{1}	10^{0}	10^{-6}	REL. F
			10^{-12}		467	327	10^{4}	10^{1}	10^{-1}	10^{-6}	REL. F
11.[0]	3	10	10^{-8}		4	3+	5.66	10^{-1}	10^{-6}	10^{-2}	REL. F
			10^{-12}		327	267+	55.9	10^{-13}	10^{-13}	10^{-26}	ABS. F
12.[0]	3	10	10^{-8}		43	34+	10.1	10^{-9}	10^{-10}	10^{-18}	ABS. F
			10^{-12}		45	36+	10.1	10^{-14}	10^{-14}	10^{-28}	ABS. F
13.[0]	4	4	10^{-8}		62	61+	10^{-4}	10^{-8}	10^{-9}	10^{-16}	ABS. F
			10^{-12}		89	88+	10^{-6}	10^{-12}	10^{-13}	10^{-24}	ABS. F
14.[0]	4	6	10^{-8}		100	78+	2.00	10^{-8}	10^{-7}	10^{-16}	ABS. F
			10^{-12}		102	80+	2.00	10^{-12}	10^{-11}	10^{-24}	ABS. F
15.	4	11	10^{-8}		35	31+	.328	10^{-2}	10^{-8}	10^{-9}	REL. F
			10^{-12}		36	32+	.328	10^{-2}	10^{-11}	10^{-9}	REL. F
16.	4	20	10^{-8}		46	34+	17.6	10^{2}	10^{-2}	10^{-8}	REL. F
			10^{-12}		47	35+	17.6	10^{2}	10^{-5}	10^{-8}	REL. F
17.	5	33	10^{-8}		69	55+	2.46	10^{-2}	10^{-5}	10^{-11}	REL. F
			10^{-12}		72	58+	2.46	10^{-2}	10^{-9}	10^{-11}	REL. F
18.[0]	6	13	10^{-8}		45	41+	18.7	10^{-1}	10^{-7}	10^{-2}	REL. F
			10^{-12}		47	43+	18.7	10^{-1}	10^{-10}	10^{-2}	REL. F
19.	11	65	10^{-8}		69	58+	9.38	10^{-1}	10^{-6}	10^{-8}	REL. F
			10^{-12}		72	61+	9.38	10^{-1}	10^{-8}	10^{-8}	REL. F

Numerical Results for DMNG

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|\|x^*\|\|_2$	$\|\|f^*\|\|_2$	$\|\|g^*\|\|_2$	est. err.	conv.
20a.	6	31	10^{-8}		35	32+	2.44	10^{-2}	10^{-6}	10^{-10}	REL. F
			10^{-12}		37	34+	2.44	10^{-2}	10^{-8}	10^{-10}	REL. F
20b.	9	31	10^{-8}		76	69+	6.06	10^{-3}	10^{-10}	10^{-13}	REL. F
			10^{-12}		79	74	6.06	10^{-3}	10^{-10}	10^{-13}	REL. F
20c.	12	31	10^{-8}		89	85+	1.28	10^{-4}	10^{-9}	10^{-8}	REL. F
			10^{-12}		148	143	2.87	10^{-5}	10^{-13}	10^{-9}	REL. F
20d.	20	31	10^{-8}		110	107+	1.06	10^{-6}	10^{-11}	10^{-12}	X
			10^{-12}		134	119	1.06	10^{-7}	10^{-12}	10^{-13}	FALSE
21a.^{0}	10	10	10^{-8}		120	98+	3.16	10^{-8}	10^{-7}	10^{-17}	ABS. F
			10^{-12}		125	103+	3.16	10^{-13}	10^{-12}	10^{-27}	X
21b.^{0}	20	20	10^{-8}		189	148+	4.47	10^{-9}	10^{-8}	10^{-17}	X
			10^{-12}		193	152+	4.47	10^{-13}	10^{-12}	10^{-25}	ABS. F
22a.^{0}	12	12	10^{-8}		143	131+	10^{-4}	10^{-8}	10^{-11}	10^{-16}	ABS. F
			10^{-12}		235	219+	10^{-6}	10^{-12}	10^{-17}	10^{-24}	ABS. F
22b.^{0}	20	20	10^{-8}		187	153+	10^{-4}	10^{-8}	10^{-8}	10^{-16}	ABS. F
			10^{-12}		344	311+	10^{-6}	10^{-12}	10^{-17}	10^{-24}	ABS. F
23a.	4	5	10^{-8}		77	57+	.500	10^{-3}	10^{-10}	10^{-10}	REL. F
			10^{-12}		78	58+	.500	10^{-3}	10^{-11}	10^{-10}	REL. F
23b.	10	11	10^{-8}		80	67+	.500	10^{-2}	10^{-9}	10^{-11}	REL. F
			10^{-12}		81	68+	.500	10^{-2}	10^{-11}	10^{-11}	REL. F
24a.	4	8	10^{-8}		364	270+	.769	10^{-3}	10^{-7}	10^{-9}	REL. F
			10^{-12}		472	355+	.759	10^{-3}	10^{-11}	10^{-11}	REL. F
24b.	10	20	10^{-8}		475	367+	.606	10^{-2}	10^{-5}	10^{-8}	REL. F
			10^{-12}		632	510+	.598	10^{-2}	10^{-9}	10^{-9}	REL. F
25a.^{0}	10	12	10^{-8}		20	19+	3.16	10^{-12}	10^{-10}	10^{-23}	ABS. F
			10^{-12}		21	20+	3.16	0.00	0.00	0.00	X
25b.^{0}	20	22	10^{-8}		25	24+	4.47	10^{-9}	10^{-7}	10^{-17}	ABS. F
			10^{-12}		26	25+	4.47	10^{-15}	10^{-13}	10^{-20}	ABS. F
26a.^{0}	10	10	10^{-8}		34	32+	.328	10^{-2}	10^{-8}	10^{-5}	REL. F
			10^{-12}		37	35+	.328	10^{-2}	10^{-10}	10^{-5}	REL. F
26b.^{0}	20	20	10^{-8}		62	59+	.231	10^{-3}	10^{-8}	10^{-5}	REL. F
			10^{-12}		65	62	.231	10^{-3}	10^{-10}	10^{-5}	REL. F
27a.^{0}	10	10	10^{-8}		13	11+	3.16	10^{-8}	10^{-7}	10^{-16}	ABS. F
			10^{-12}		16	14+	3.16	10^{-12}	10^{-12}	10^{-24}	ABS. F
27b.^{0}	20	20	10^{-8}		15	12+	4.47	10^{-8}	10^{-7}	10^{-16}	ABS. F
			10^{-12}		18	15+	4.47	10^{-13}	10^{-12}	10^{-27}	ABS. F
28a.^{0}	10	10	10^{-8}		31	25+	.412	10^{-9}	10^{-9}	10^{-18}	ABS. F
			10^{-12}		34	28+	.412	10^{-14}	10^{-14}	10^{-28}	ABS. F
28b.^{0}	20	20	10^{-8}		60	48+	.571	10^{-8}	10^{-8}	10^{-16}	ABS. F
			10^{-12}		64	52+	.571	10^{-13}	10^{-13}	10^{-26}	ABS. F

Numerical Results for DMNG

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv
29a.[0]	10	10	10^{-8}		8	7+	.412	10^{-9}	10^{-9}	10^{-18}	ABS. F
			10^{-12}		10	9+	.412	10^{-13}	10^{-13}	10^{-25}	ABS. F
29b.[0]	20	20	10^{-8}		8	7+	.571	10^{-9}	10^{-9}	10^{-18}	ABS. F
			10^{-12}		10	9+	.571	10^{-13}	10^{-13}	10^{-25}	ABS. F
30a.[0]	10	10	10^{-8}		51	39+	2.05	10^{-8}	10^{-7}	10^{-15}	X
			10^{-12}		57	45+	2.05	10^{-13}	10^{-12}	10^{-28}	X
30b.[0]	20	20	10^{-8}		65	45+	3.04	10^{-8}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		88	61+	3.04	10^{-12}	10^{-11}	10^{-24}	ABS. F
31a.[0]	10	10	10^{-8}		46	26+	1.80	10^{-8}	10^{-8}	10^{-16}	ABS. F
			10^{-12}		60	35+	1.80	10^{-12}	10^{-12}	10^{-28}	ABS. F
31b.[0]	20	20	10^{-8}		47	26+	2.66	10^{-8}	10^{-7}	10^{-16}	ABS. F
			10^{-12}		63	36	2.66	10^{-12}	10^{-11}	10^{-23}	X
32.[L]	10	20	10^{-8}		6	4	3.16	10^{0}	10^{-16}	0.00	X, REL. F
			10^{-12}		6	4	3.16	10^{0}	10^{-16}	0.00	X, REL. F
33.[L]	10	20	10^{-8}		4	2	1.46	10^{0}	10^{-8}	10^{-6}	X, REL. F
			10^{-12}		4	2	1.46	10^{0}	10^{-8}	10^{-6}	X, REL. F
34.[L]	10	20	10^{-8}		5	3	1.78	10^{0}	10^{-12}	10^{-6}	X, REL. F
			10^{-12}		5	3	1.78	10^{0}	10^{-12}	10^{-6}	X, REL. F
35a.	8	8	10^{-8}		34	24	1.65	10^{-1}	10^{-5}	10^{-9}	REL. F
			10^{-12}		38	27+	1.65	10^{-1}	10^{-8}	10^{-9}	REL. F
35b.[0]	9	9	10^{-8}		44	32+	1.73	10^{-9}	10^{-9}	10^{-18}	ABS. F
			10^{-12}		46	34+	1.73	10^{-12}	10^{-12}	10^{-24}	ABS. F
35c.	10	10	10^{-8}		41	31+	1.81	10^{-1}	10^{-6}	10^{-3}	REL. F
			10^{-12}		45	36+	1.81	10^{-1}	10^{-8}	10^{-3}	REL. F
36a.[0]	4	4	10^{-8}		(4000)	(2891)	17.0	10^{-6}	10^{-6}	10^{-11}	F LIM.
			10^{-12}		(4000)	(2891)	17.0	10^{-6}	10^{-6}	10^{-11}	F LIM.
36b.[0]	9	9	10^{-8}		(9000)	(6426)	228.	10^{-6}	10^{-7}	10^{-12}	F LIM.
			10^{-12}		(9000)	(6426)	228.	10^{-6}	10^{-7}	10^{-12}	F LIM.
36c.[0]	9	9	10^{-8}		69	64+	1.73	10^{-8}	10^{-9}	10^{-16}	ABS. F
			10^{-12}		101	96+	1.73	10^{-12}	10^{-13}	10^{-24}	ABS. F
36d.[0]	9	9	10^{-8}		(9000)	(6486)	228.	10^{-6}	10^{-7}	10^{-12}	F LIM.
			10^{-12}		(9000)	(6486)	228.	10^{-6}	10^{-7}	10^{-12}	F LIM.
37.	2	16	10^{-8}		22	10+	9.05	10^{1}	10^{-6}	10^{-1}	REL. F
			10^{-12}		22	10+	9.05	10^{1}	10^{-6}	10^{-1}	REL. F
38.	3	16	10^{-8}		31	17+	26.1	10^{1}	10^{-4}	10^{-6}	REL. F
			10^{-12}		32	18+	26.1	10^{1}	10^{-5}	10^{-6}	REL. F

Numerical Results for DMNG

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
39a.	2	3	10^{-8}		9	8+	10^{-6}	10^{-1}	10^{-6}	10^{-7}	REL. F
			10^{-12}		11	10+	10^{-6}	10^{-1}	10^{-10}	10^{-7}	REL. F
39b.	2	3	10^{-8}		9	8+	10^{-7}	10^{-1}	10^{-7}	10^{-7}	REL. F
			10^{-12}		10	9+	10^{-7}	10^{-1}	10^{-9}	10^{-7}	REL. F
39c.	2	3	10^{-8}		6	5+	10^{-7}	10^{-1}	10^{-7}	10^{-7}	REL. F
			10^{-12}		7	6+	10^{-7}	10^{-1}	10^{-12}	10^{-7}	REL. F
39d.	2	3	10^{-8}		8	6+	10^{-8}	10^{-1}	10^{-8}	10^{-7}	REL. F
			10^{-12}		9	7+	10^{-7}	10^{-1}	10^{-9}	10^{-7}	REL. F
39e.	2	3	10^{-8}		11	8+	10^{-7}	10^{-1}	10^{-6}	10^{-7}	REL. F
			10^{-12}		12	9+	10^{-8}	10^{-1}	10^{-8}	10^{-7}	REL. F
39f.	2	3	10^{-8}		11	8+	10^{-9}	10^{-1}	10^{-8}	10^{-7}	REL. F
			10^{-12}		11	8+	10^{-9}	10^{-1}	10^{-8}	10^{-7}	REL. F
39g.	2	3	10^{-8}		17	13+	10^{-9}	10^{-1}	10^{-6}	10^{-7}	REL. F
			10^{-12}		18	14+	10^{-10}	10^{-1}	10^{-8}	10^{-7}	REL. F
40a.	3	4	10^{-8}		11	10+	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
			10^{-12}		12	11+	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
40b.	3	4	10^{-8}		10	9+	10^{-6}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		12	11+	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
40c.	3	4	10^{-8}		9	7+	10^{-7}	10^{0}	10^{-7}	10^{-7}	REL. F
			10^{-12}		10	8+	10^{-7}	10^{0}	10^{-11}	10^{-7}	REL. F
40d.	3	4	10^{-8}		10	7+	10^{-7}	10^{0}	10^{-7}	10^{-7}	REL. F
			10^{-12}		11	8+	10^{-7}	10^{0}	10^{-9}	10^{-7}	REL. F
40e.	3	4	10^{-8}		11	8+	10^{-7}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		13	10+	10^{-7}	10^{0}	10^{-9}	10^{-7}	REL. F
40f.	3	4	10^{-8}		14	11+	10^{-7}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		16	13+	10^{-8}	10^{0}	10^{-8}	10^{-7}	REL. F
40g.	3	4	10^{-8}		18	14+	10^{-8}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		20	16+	10^{-9}	10^{0}	10^{-6}	10^{-7}	REL. F
41a.	5	10	10^{-8}		11	8+	10^{-6}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		13	10+	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
41b.	5	10	10^{-8}		11	8+	10^{-6}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		13	10+	10^{-6}	10^{0}	10^{-9}	10^{-7}	REL. F
41c.	5	10	10^{-8}		13	9+	10^{-6}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		14	10+	10^{-6}	10^{0}	10^{-7}	10^{-7}	REL. F
41d.	5	10	10^{-8}		17	12+	10^{-5}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		20	15+	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
41e.	5	10	10^{-8}		24	20+	10^{-6}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		26	22+	10^{-7}	10^{0}	10^{-7}	10^{-7}	REL. F
41f.	5	10	10^{-8}		27	22+	10^{-6}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		31	26+	10^{-7}	10^{0}	10^{-6}	10^{-7}	REL. F
41g.	5	10	10^{-8}		32	26+	10^{-6}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		35	29+	10^{-8}	10^{0}	10^{-8}	10^{-7}	REL. F

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
42a.^{0}	4	24	10^{-8}		53	41+	60.8	10^{-8}	10^{-6}	10^{-16}	X
			10^{-12}		56	43+	60.8	10^{-13}	10^{-11}	10^{-25}	X
42b.^{0}	4	24	10^{-8}	0.9	103	69+	61.9	10^{-10}	10^{-8}	10^{-20}	X
			10^{-12}	0.9	104	70+	61.9	10^{-13}	10^{-10}	10^{-26}	X
42c.^{0}	4	24	10^{-8}		76	52+	60.3	10^{-9}	10^{-6}	10^{-17}	X
			10^{-12}		78	54+	60.3	10^{-13}	10^{-11}	10^{-26}	X
42d.^{0}	4	24	10^{-8}		61	48+	60.3	10^{-7}	10^{-5}	10^{-14}	X
			10^{-12}		64	51+	60.3	10^{-12}	10^{-10}	10^{-24}	X
43a.^{0}	5	16	10^{-8}		49	34+	54.0	10^{-9}	10^{-7}	10^{-18}	X
			10^{-12}		51	36+	54.0	10^{-12}	10^{-10}	10^{-24}	X
43b.^{0}	5	16	10^{-8}		58	37+	54.0	10^{-9}	10^{-7}	10^{-16}	X
			10^{-12}		60	39+	54.0	10^{-14}	10^{-12}	10^{-27}	X
43c.^{0}	5	16	10^{-8}		41	29+	54.0	10^{-9}	10^{-6}	10^{-17}	X
			10^{-12}		44	32+	54.0	10^{-13}	10^{-12}	10^{-26}	X
43d.^{0}	5	16	10^{-8}		57	44+	54.0	10^{-9}	10^{-7}	10^{-17}	ABS. F
			10^{-12}		60	47+	54.0	10^{-13}	10^{-11}	10^{-26}	X
43e.^{0}	5	16	10^{-8}		51	41+	54.0	10^{-9}	10^{-8}	10^{-18}	X
			10^{-12}		53	43+	54.0	10^{-14}	10^{-11}	10^{-27}	X
43f.^{0}	5	16	10^{-8}		45	36+	54.0	10^{-8}	10^{-6}	10^{-16}	X
			10^{-12}		48	39+	54.0	10^{-13}	10^{-10}	10^{-25}	X
44a.^{0}	6	6	10^{-8}		441	341+	4.03	10^{-9}	10^{-8}	10^{-19}	ABS. F
			10^{-12}		444	344+	4.03	10^{-12}	10^{-11}	10^{-25}	ABS. F
44b.^{0}	6	6	10^{-8}		31	24+	3.52	10^{-9}	10^{-8}	10^{-18}	X
			10^{-12}		34	27+	3.52	10^{-14}	10^{-12}	10^{-26}	X
44c.^{0}	6	6	10^{-8}		3726	2748+	20.6	10^{-8}	10^{-5}	10^{-16}	ABS. F
			10^{-12}		3731	2753+	20.6	10^{-12}	10^{-10}	10^{-25}	ABS. F
44d.^{0}	6	6	10^{-8}		2752	1928+	15.3	10^{-1}	10^{0}	10^{-2}	REL. F
			10^{-12}		3865	2915+	15.3	10^{-12}	10^{-9}	10^{-24}	X
44e.^{0}	6	6	10^{-8}		2104	1550+	12.3	10^{-1}	10^{0}	10^{-2}	REL. F
			10^{-12}		2813	2098+	9.27	10^{-13}	10^{-10}	10^{-25}	ABS. F
45a.^{0}	8	8	10^{-8}		284	227+	4.06	10^{-8}	10^{-7}	10^{-16}	ABS. F
			10^{-12}		288	231+	4.06	10^{-12}	10^{-11}	10^{-25}	ABS. F
45b.^{0}	8	8	10^{-8}		36	28+	3.56	10^{-8}	10^{-7}	10^{-17}	ABS. F
			10^{-12}		40	32+	3.56	10^{-13}	10^{-12}	10^{-26}	X
45c.^{0}	8	8	10^{-8}		6197	4538+	20.6	10^{-8}	10^{-5}	10^{-16}	ABS. F
			10^{-12}		6200	4541+	20.6	10^{-12}	10^{-9}	10^{-24}	X
45d.^{0}	8	8	10^{-8}		7929	5976+	15.3	10^{-8}	10^{-5}	10^{-16}	ABS. F
			10^{-12}		7934	5981+	15.3	10^{-14}	10^{-11}	10^{-28}	ABS. F
45e.^{0}	8	8	10^{-8}		3341	2511+	9.31	10^{-9}	10^{-7}	10^{-17}	ABS. F
			10^{-12}		3346	2517+	9.31	10^{-13}	10^{-11}	10^{-25}	ABS. F

Numerical Results for DMNH

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
1.^{0}	2	2	10^{-8}		32	24+	1.41	10^{-14}	10^{-13}	10^{-28}	ABS. F
			10^{-12}		32	24+	1.41	10^{-14}	10^{-13}	10^{-28}	ABS. F
2.^{0}	2	2	10^{-8}		10	7+	11.4	10^{1}	10^{-11}	10^{1}	REL. F
			10^{-12}		10	7+	11.4	10^{1}	10^{-11}	10^{1}	REL. F
3.^{0}	2	2	10^{-8}		130	101+	9.11	10^{-9}	10^{-4}	10^{-17}	ABS. F
			10^{-12}		132	103+	9.11	10^{-16}	10^{-11}	10^{-31}	ABS. F
4.^{0}	2	3	10^{-8}		22	10+	10^{6}	10^{-11}	10^{-5}	10^{-22}	X
			10^{-12}		23	11+	10^{6}	10^{-22}	10^{-22}	10^{-44}	X
5.^{0}	2	3	10^{-8}		11	8+	3.04	10^{-10}	10^{-10}	10^{-20}	ABS. F
			10^{-12}		12	9+	3.04	$10^{0.00}$	$10^{0.00}$	$10^{0.00}$	ABS. F
6.	2	10	10^{-8}		11	10+	.365	10^{1}	10^{-9}	10^{-6}	REL. F
			10^{-12}		11	10+	.365	10^{1}	10^{-9}	10^{-6}	REL. F
7.^{0}	3	3	10^{-8}		16	12+	1.00	10^{-8}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		17	13+	1.00	10^{-16}	10^{-16}	10^{-33}	ABS. F
8.	3	15	10^{-8}		9	8+	2.60	10^{-1}	10^{-12}	10^{-8}	REL. F
			10^{-12}		10	9	2.60	10^{-1}	10^{-12}	10^{-8}	REL. F
9.	3	15	10^{-8}		4	3+	1.08	10^{-4}	10^{-16}	10^{-14}	X, REL. F
			10^{-12}		4	3+	1.08	10^{-4}	10^{-16}	10^{-14}	REL. F
10.	3	16	10^{-8}		387	244+	10^{4}	10^{1}	10^{-1}	10^{-6}	REL. F
			10^{-12}		388	245	10^{4}	10^{1}	10^{-1}	10^{-6}	REL. F
11.^{0}	3	10	10^{-8}		290	167+	55.9	10^{-9}	10^{-9}	10^{-18}	ABS. F
			10^{-12}		292	169+	55.9	10^{-16}	10^{-16}	[illegible]	ABS. F
12.^{0}	3	10	10^{-8}		24	19+	10.1	10^{-13}	10^{-13}	10^{-26}	ABS. F
			10^{-12}		24	19+	10.1	10^{-13}	10^{-13}	10^{-26}	ABS. F
13.^{0}	4	4	10^{-8}		27	26+	10^{-4}	10^{-8}	10^{-12}	10^{-16}	ABS. F
			10^{-12}		38	37+	10^{-6}	10^{-12}	10^{-17}	10^{-24}	ABS. F
14.^{0}	4	6	10^{-8}		42	32+	2.00	10^{-12}	10^{-11}	10^{-24}	ABS. F
			10^{-12}		49	38+	2.00	0.00	0.00	0.00	ABS. F
15.	4	11	10^{-8}		11	8+	.328	10^{-2}	10^{-13}	10^{-9}	REL. F
			10^{-12}		12	9	.328	10^{-2}	10^{-13}	10^{-9}	REL. F
16.	4	20	10^{-8}		11	9+	17.6	10^{2}	10^{-6}	10^{-8}	REL. F
			10^{-12}		13	11	17.6	10^{2}	10^{-11}	10^{-8}	X, REL. F
17.	5	33	10^{-8}	0.2	46	32+	2.46	10^{-2}	10^{-7}	10^{-11}	REL. F
			10^{-12}	0.2	47	33+	2.46	10^{-2}	10^{-12}	10^{-11}	REL. F
18.^{0}	6	13	10^{-8}		(6000)	(1824+)	283.	10^{-1}	10^{-5}	10^{-1}	F LIM.
			10^{-12}		(6000)	(1820+)	275.	10^{-1}	10^{-9}	10^{-1}	F LIM.
19.	11	65	10^{-8}		23	17+	9.38	10^{-1}	10^{-9}	10^{-8}	REL. F
			10^{-12}		24	18+	9.38	10^{-1}	10^{-10}	10^{-8}	REL. F

Numerical Results for DNNH

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
20a.	6	31	10^{-8}		15	14	2.44	10^{-2}	10^{-15}	10^{-10}	X. REL F
			10^{-12}		15	14	2.44	10^{-2}	10^{-15}	10^{-10}	X. REL F
20b.	9	31	10^{-8}		20	16+	6.06	10^{-3}	10^{-13}	10^{-13}	REL F
			10^{-12}		22	18	6.06	10^{-3}	10^{-14}	10^{-13}	X. REL. F
20c.	12	31	10^{-8}		24	19	16.6	10^{-5}	10^{-13}	10^{-16}	X. REL. F
			10^{-12}		24	19	16.6	10^{-5}	10^{-13}	10^{-16}	X. REL. F
20d.	20	31	10^{-8}		50	26+	1.10	10^{-8}	10^{-13}	10^{-16}	ABS. F
			10^{-12}		(149)	(55+)	1.16	10^{-8}	10^{-13}	10^{-16}	LOOP
21a.^{0}	10	10	10^{-8}		25	22+	3.16	10^{-9}	10^{-8}	10^{-19}	ABS. F
			10^{-12}		26	23+	3.16	10^{-16}	10^{-14}	10^{-30}	ABS. F
21b.^{0}	20	20	10^{-8}		27	23+	4.47	10^{-14}	10^{-13}	10^{-28}	ABS. F
			10^{-12}		27	23+	4.47	10^{-14}	10^{-13}	10^{-28}	ABS. F
22a.^{0}	12	12	10^{-8}		28	27+	10^{-4}	10^{-8}	10^{-12}	10^{-16}	ABS. F
			10^{-12}		40	39+	10^{-6}	10^{-12}	10^{-18}	10^{-24}	ABS. F
22b.^{0}	20	20	10^{-8}		29	28+	10^{-4}	10^{-8}	10^{-12}	10^{-17}	ABS. F
			10^{-12}		40	39+	10^{-6}	10^{-12}	10^{-18}	10^{-24}	ABS. F
23a.	4	5	10^{-8}		42	36+	.500	10^{-3}	10^{-12}	10^{-10}	REL. F
			10^{-12}		43	37	.500	10^{-3}	10^{-12}	10^{-10}	REL. F
23b.	10	11	10^{-8}		44	37+	.500	10^{-2}	10^{-9}	10^{-11}	REL. F
			10^{-12}		45	38+	.500	10^{-2}	10^{-14}	10^{-11}	REL. F
24a.	4	8	10^{-8}		126	110+	.759	10^{-3}	10^{-7}	10^{-11}	REL. F
			10^{-12}		128	112+	.759	10^{-3}	10^{-13}	10^{-11}	REL. F
24b.	10	20	10^{-8}		158	106+	.598	10^{-2}	10^{-7}	10^{-9}	REL. F
			10^{-12}		162	110	.598	10^{-2}	10^{-16}	10^{-9}	X. REL. F
25a.^{0}	10	12	10^{-8}		15	14+	3.16	10^{-13}	10^{-12}	10^{-26}	ABS. F
			10^{-12}		15	14+	3.16	10^{-13}	10^{-12}	10^{-26}	ABS. F
25b.^{0}	20	22	10^{-8}		18	17+	4.47	10^{-10}	10^{-8}	10^{-20}	ABS. F
			10^{-12}		19	18	4.47	10^{-15}	10^{-13}	10^{-30}	X
26a.^{0}	10	10	10^{-8}		11	9+	.328	10^{-2}	10^{-11}	10^{-5}	REL. F
			10^{-12}		12	10	.328	10^{-2}	10^{-11}	10^{-5}	REL. F
26b.^{0}	20	20	10^{-8}		20	16	.228	10^{-3}	10^{-11}	10^{-6}	X. REL. F
			10^{-12}		20	16	.228	10^{-3}	10^{-11}	10^{-6}	X. REL. F
27a.^{0}	10	10	10^{-8}		9	7+	3.16	10^{-10}	10^{-10}	10^{-21}	ABS. F
			10^{-12}		10	8+	3.16	10^{-15}	10^{-14}	10^{-29}	ABS. F
27b.^{0}	20	20	10^{-8}		11	9+	4.47	10^{-9}	10^{-9}	10^{-18}	ABS. F
			10^{-12}		12	11+	4.47	10^{-14}	10^{-12}	10^{-27}	ABS. F
28a.^{0}	10	10	10^{-8}		4	3+	.412	10^{-12}	10^{-13}	10^{-24}	ABS. F
			10^{-12}		4	3+	.412	10^{-12}	10^{-13}	10^{-24}	ABS. F
28b.^{0}	20	20	10^{-8}		4	3+	.571	10^{-13}	10^{-14}	10^{-25}	ABS. F
			10^{-12}		4	3+	.571	10^{-13}	10^{-14}	10^{-25}	ABS. F

Numerical Results for DMNH

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
29a.^{0}	10	10	10^{-8}		4	3+	.412	10^{-10}	10^{-10}	10^{-20}	ABS. F
			10^{-12}		5	4+	.412	10^{-16}	10^{-16}	10^{-32}	ABS. F
29b.^{0}	20	20	10^{-8}		4	3+	.571	10^{-10}	10^{-10}	10^{-20}	ABS. F
			10^{-12}		5	4+	.571	10^{-16}	10^{-16}	10^{-32}	ABS. F
30a.^{0}	10	10	10^{-8}		6	5+	2.05	10^{-8}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		7	6+	2.05	10^{-15}	10^{-15}	10^{-31}	ABS. F
30b.^{0}	20	20	10^{-8}		6	5+	3.04	10^{-8}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		7	6+	3.04	10^{-15}	10^{-15}	10^{-30}	ABS. F
31a.^{0}	10	10	10^{-8}		9	8+	1.80	10^{-13}	10^{-12}	10^{-26}	ABS. F
			10^{-12}		9	8+	1.80	10^{-13}	10^{-12}	10^{-26}	ABS. F
31b.^{0}	20	20	10^{-8}		9	8+	1.80	10^{-13}	10^{-12}	10^{-26}	ABS. F
			10^{-12}		9	8+	1.80	10^{-13}	10^{-12}	10^{-26}	ABS. F
32.^{L}	10	20	10^{-8}		6	4	3.16	10^{0}	10^{-32}	0.00	X. REL. F
			10^{-12}		6	4	3.16	10^{0}	10^{-32}	0.00	X. REL. F
33.^{L}	10	20	10^{-8}		5	4	1.46	10^{0}	10^{-11}	10^{-6}	SING.
			10^{-12}		5	4	1.46	10^{0}	10^{-11}	10^{-6}	SING.
34.^{L}	10	20	10^{-8}		6	5	1.78	10^{0}	10^{-11}	10^{-6}	SING.
			10^{-12}		6	5	1.78	10^{0}	10^{-11}	10^{-6}	SING.
35a.	8	8	10^{-8}		14	11	1.65	10^{-1}	10^{-9}	10^{-9}	X. REL. F
			10^{-12}		14	11	1.65	10^{-1}	10^{-9}	10^{-9}	REL. F
35b.^{0}	9	9	10^{-8}		17	12+	1.73	10^{-9}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		18	13+	1.73	10^{-15}	10^{-14}	10^{-29}	ABS. F
35c.	10	10	10^{-8}		19	11+	1.76	10^{-1}	10^{-9}	10^{-9}	REL. F
			10^{-12}		20	12	1.76	10^{-1}	10^{-9}	10^{-9}	REL. F
36a.^{0}	4	4	10^{-8}		(4000)	(2190)	17.8	10^{-6}	10^{-6}	10^{-12}	F LIM.
			10^{-12}		(4000)	(2190)	17.8	10^{-6}	10^{-6}	10^{-12}	F LIM.
36b.^{0}	9	9	10^{-8}		(1832)	(611)	9.15	10^{-5}	10^{-9}	10^{-9}	TIME
			10^{-12}		(1881)	(625)	9.24	10^{-5}	10^{-5}	10^{-9}	TIME
36c.^{0}	9	9	10^{-8}		31	30+	1.73	10^{-8}	10^{-10}	10^{-17}	ABS. F
			10^{-12}		35	33+	1.73	10^{-13}	10^{-16}	10^{-25}	ABS. F
36d.^{0}	9	9	10^{-8}		(1900)	(638)	9.38	10^{-5}	10^{-5}	10^{-9}	TIME
			10^{-12}		(1859)	(627)	9.32	10^{-5}	10^{-5}	10^{-9}	TIME
37.	2	16	10^{-8}		16	7+	9.05	10^{1}	10^{-7}	10^{-1}	REL. F
			10^{-12}		17	8	9.05	10^{1}	10^{-7}	10^{-1}	REL. F
38.	3	16	10^{-8}		14	8+	26.1	10^{1}	10^{-10}	10^{-6}	REL. F
			10^{-12}		14	8+	26.1	10^{1}	10^{-10}	10^{-6}	REL. F

Numerical Results for DMNH

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
39a.	2	3	10^{-8}		4	3+	10^{-6}	10^{-1}	10^{-17}	10^{-7}	REL. F
			10^{-12}		4	3+	10^{-6}	10^{-1}	10^{-17}	10^{-7}	REL. F
39b.	2	3	10^{-8}		4	3+	10^{-7}	10^{-1}	10^{-11}	10^{-7}	REL. F
			10^{-12}		4	3+	10^{-7}	10^{-1}	10^{-11}	10^{-7}	REL. F
39c.	2	3	10^{-8}		4	3+	10^{-7}	10^{-1}	10^{-13}	10^{-7}	REL. F
			10^{-12}		5	4	10^{-7}	10^{-1}	10^{-13}	10^{-7}	REL. F
39d.	2	3	10^{-8}		6	5+	10^{-7}	10^{-1}	10^{-8}	10^{-7}	REL. F
			10^{-12}		6	5+	10^{-7}	10^{-1}	10^{-8}	10^{-7}	REL. F
39e.	2	3	10^{-8}		8	7+	10^{-8}	10^{-1}	10^{-14}	10^{-7}	REL. F
			10^{-12}		8	7+	10^{-8}	10^{-1}	10^{-14}	10^{-7}	REL. F
39f.	2	3	10^{-8}		11	10	10^{-9}	10^{-1}	10^{-8}	10^{-7}	REL. F
			10^{-12}		11	10	10^{-9}	10^{-1}	10^{-8}	10^{-7}	REL. F
39g.	2	3	10^{-8}		13	12+	10^{-10}	10^{-1}	10^{-9}	10^{-7}	REL. F
			10^{-12}		14	13	10^{-10}	10^{-1}	10^{-9}	10^{-7}	REL. F
40a.	2	3	10^{-8}		4	3+	10^{-6}	10^{-1}	10^{-15}	10^{-7}	REL. F
			10^{-12}		4	3+	10^{-6}	10^{-1}	10^{-15}	10^{-7}	REL. F
40b.	3	4	10^{-8}		4	3+	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
			10^{-12}		5	4	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
40c.	3	4	10^{-8}		5	4+	10^{-7}	10^{0}	10^{-12}	10^{-7}	REL. F
			10^{-12}		5	4+	10^{-7}	10^{0}	10^{-12}	10^{-7}	REL. F
40d.	3	4	10^{-8}		5	4+	10^{-7}	10^{0}	10^{-8}	10^{-7}	REL. F
			10^{-12}		6	5+	10^{-7}	10^{0}	10^{-16}	10^{-7}	REL. F
40e.	3	4	10^{-8}		7	6+	10^{-7}	10^{0}	10^{-8}	10^{-7}	REL. F
			10^{-12}		7	6+	10^{-7}	10^{0}	10^{-8}	10^{-7}	REL. F
40f.	3	4	10^{-8}		10	9+	10^{-8}	10^{0}	10^{-11}	10^{-7}	REL. F
			10^{-12}		10	9+	10^{-8}	10^{0}	10^{-11}	10^{-7}	REL. F
40g.	3	4	10^{-8}		13	12+	10^{-9}	10^{0}	10^{-15}	10^{-7}	REL. F
			10^{-12}		13	12+	10^{-9}	10^{0}	10^{-15}	10^{-7}	REL. F
41a.	5	10	10^{-8}		4	3	10^{-6}	10^{0}	10^{-9}	10^{-7}	REL. F
			10^{-12}		4	3	10^{-6}	10^{0}	10^{-9}	10^{-7}	REL. F
41b.	5	10	10^{-8}		4	3+	10^{-6}	10^{0}	10^{-13}	10^{-7}	REL. F
			10^{-12}		4	3+	10^{-6}	10^{0}	10^{-13}	10^{-7}	REL. F
41c.	5	10	10^{-8}		8	7+	10^{-6}	10^{0}	10^{-12}	10^{-7}	REL. F
			10^{-12}		8	7+	10^{-6}	10^{0}	10^{-12}	10^{-7}	REL. F
41d.	5	10	10^{-8}		8	7+	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
			10^{-12}		9	8+	10^{-6}	10^{0}	10^{-14}	10^{-7}	REL. F
41e.	5	10	10^{-8}		11	10+	10^{-7}	10^{0}	10^{-10}	10^{-7}	REL. F
			10^{-12}		12	11	10^{-7}	10^{0}	10^{-10}	10^{-7}	REL. F
41f.	5	10	10^{-8}		14	13	10^{-7}	10^{0}	10^{-11}	10^{-7}	REL. F
			10^{-12}		14	13	10^{-7}	10^{0}	10^{-11}	10^{-7}	REL. F
41g.	5	10	10^{-8}		17	16+	10^{-8}	10^{0}	10^{-12}	10^{-7}	REL. F
			10^{-12}		17	16+	10^{-8}	10^{0}	10^{-12}	10^{-7}	REL. F

Numerical Results for DMNH

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
42a.^{0}	4	24	10^{-8}		28	20+	60.8	10^{-12}	10^{-10}	10^{-25}	ABS. F
			10^{-12}		28	20+	60.8	10^{-12}	10^{-10}	10^{-25}	ABS. F
42b.^{0}	4	24	10^{-8}		35	26+	61.9	10^{-10}	10^{-8}	10^{-21}	ABS. F
			10^{-12}		36	27+	61.9	10^{-12}	10^{-10}	10^{-25}	X
42c.^{0}	4	24	10^{-8}		30	22+	60.3	10^{-12}	10^{-10}	10^{-23}	ABS. F
			10^{-12}		31	23+	60.3	10^{-13}	10^{-10}	10^{-26}	X
42d.^{0}	4	24	10^{-8}		30	21+	60.3	10^{-14}	10^{-11}	10^{-27}	X
			10^{-12}		30	21+	60.3	10^{-14}	10^{-11}	10^{-27}	ABS. F
43a.^{0}	5	16	10^{-8}		22	16+	54.0	10^{-14}	10^{-11}	10^{-27}	X
			10^{-12}		22	16+	54.0	10^{-14}	10^{-11}	10^{-27}	ABS. F
43b.^{0}	5	16	10^{-8}		26	20+	54.0	10^{-9}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		27	21+	54.0	10^{-14}	10^{-12}	10^{-28}	ABS. F
43c.^{0}	5	16	10^{-8}		21	17+	54.0	10^{-13}	10^{-11}	10^{-26}	X
			10^{-12}		21	17+	54.0	10^{-13}	10^{-11}	10^{-26}	ABS. F
43d.^{0}	5	16	10^{-8}	0.9	27	18+	54.0	10^{-9}	10^{-7}	10^{-18}	ABS. F
			10^{-12}	0.9	28	19+	54.0	10^{-14}	10^{-11}	10^{-27}	ABS. F
43e.^{0}	5	16	10^{-8}	0.9	28	20+	54.0	10^{-10}	10^{-8}	10^{-20}	ABS. F
			10^{-12}	0.9	29	21+	54.0	10^{-14}	10^{-12}	10^{-28}	ABS. F
43f.^{0}	5	16	10^{-8}		17	14+	54.0	10^{-9}	10^{-7}	10^{-18}	ABS. F
			10^{-12}		18	15+	54.0	10^{-14}	10^{-12}	10^{-28}	ABS. F
44a.^{0}	6	6	10^{-8}		179	150+	4.03	10^{-11}	10^{-9}	10^{-22}	ABS. F
			10^{-12}		180	151+	4.03	10^{-16}	10^{-15}	10^{-33}	ABS. F
44b.^{0}	6	6	10^{-8}		9	7+	3.52	10^{-10}	10^{-9}	10^{-19}	ABS. F
			10^{-12}		10	8+	3.52	10^{-15}	10^{-13}	10^{-30}	ABS. F
44c.^{0}	6	6	10^{-8}		194	179+	20.6	10^{-10}	10^{-7}	10^{-20}	ABS. F
			10^{-12}		195	180+	20.6	10^{-14}	10^{-10}	10^{-28}	ABS. F
44d.^{0}	6	6	10^{-8}		187	179+	15.3	10^{-9}	10^{-6}	10^{-17}	ABS. F
			10^{-12}		188	180+	15.3	10^{-14}	10^{-10}	10^{-27}	ABS. F
44e.^{0}	6	6	10^{-8}		219	210+	9.27	10^{-8}	10^{-6}	10^{-16}	ABS. F
			10^{-12}		220	211+	9.27	10^{-13}	10^{-10}	10^{-26}	ABS. F
45a.^{0}	8	8	10^{-8}		63	49+	4.06	10^{-8}	10^{-6}	10^{-16}	ABS. F
			10^{-12}		64	50+	4.06	10^{-14}	10^{-12}	10^{-27}	ABS. F
45b.^{0}	8	8	10^{-8}		15	11+	3.56	10^{-8}	10^{-7}	10^{-16}	ABS. F
			10^{-12}		16	12+	3.56	10^{-15}	10^{-14}	10^{-30}	ABS. F
45c.^{0}	8	8	10^{-8}		321	300+	20.6	10^{-9}	10^{-6}	10^{-18}	ABS. F
			10^{-12}		322	301+	20.6	10^{-14}	10^{-10}	10^{-28}	ABS. F
45d.^{0}	8	8	10^{-8}		328	292+	15.3	10^{-11}	10^{-7}	10^{-21}	ABS. F
			10^{-12}		329	293+	15.3	10^{-16}	10^{-13}	10^{-31}	ABS. F
45e.^{0}	8	8	10^{-8}		351	288+	9.31	10^{-11}	10^{-9}	10^{-23}	ABS. F
			10^{-12}		352	289+	9.31	10^{-15}	10^{-12}	10^{-29}	ABS. F

5.3 Quasi-Newton (BFGS) Linesearch Method (SOL/NAG NPSOL)

5.3.1 Software and Algorithm

The results were obtained using subroutine NPSOL from the Systems Optimization Laboratory (SOL), Stanford University, also available in the NAG Library. In NPSOL a search direction is determined at each iteration from a subproblem of the form

$$x \min_{p \in \Re^n} g_k^T p + \frac{1}{2} p^T H_k p.$$

where the Hessian matrix H_k is calculated using the BFGS method initialized with I (see Section 2.4.2). This is followed by a linesearch that uses both function and gradient information to obtain a steplength along the search direction [Gill et al. (1979)].

5.3.2 Parameters

Parameters were kept at their default values with the following exceptions † :

Infinite Bound Size	-	10^{20}
Infinite Step Size	-	10^{20}
Iteration Limit	-	10000†
Optimality Tolerance	-	varied; see tables
Step Limit	-	usually 2.0 (default) ‡

† In those cases (Problems 44c., d. ($n = m = 6$) and 45c., e. ($n = m = 8$) in which the iteration limit was actually reached, the results listed in the tables are taken from first iteration in which the number of function evaulations reaches or exceeds $1000n$.

‡ In some cases the default **Step Limit** = 2.0 was too large and overflow occured during function evaluation in the linesearch. These cases are indicated in the tables by giving the lower value of **Step Limit** that was subsequently used to obtain the results in the column labeled "**Step Lim**".

See Gill et al. [1986] for details concerning the parameters.

5.3.3 Convergence Criteria

The following quantities will be used in describing the convergence criteria :

objective function	:	$\mathcal{F}_k \;\; (= \frac{1}{2} f_k^T f_k)$
objective gradient	:	$g_k = \nabla \mathcal{F}_k \;\; (= J_k^T f_k)$
optimality tolerance	:	ϵ_{opt}

The sequence of iterates generated by NPSOL is judged to have converged if the following two conditions hold:

$$\alpha_k \|p_k\|_2 \le \sqrt{\epsilon_{opt}}(1 + \|x_k\|_2) \tag{5.3.1}$$

and

$$\|g_k\|_2 \le \sqrt{\epsilon_{opt}}(1 + \max\,\{(1 + |\mathcal{F}_k|)\|g_k\|_2\}) \tag{5.3.2}$$

or if

$$\|g_k\|_2 \le \epsilon_M^{0.8}(1 + \max\,\{(1 + |\mathcal{F}_k|), \|g_k\|_2\}). \tag{5.3.3}$$

Condition (5.3.1) is meant to ensure that the sequence $\{x_k\}$ has converged, while conditions (5.3.2) and (5.3.3) are intended to test whether the requirement that the gradient vanish is approximately satisfied at x_k. Condition (5.3.3) allows NPSOL to accept a point as a local mimimum if a more restrictive test on the necessary condition than (5.3.2) is satisfied, but condition (5.3.1) does not hold. For a detailed discussion of convergence criteria similar to these, see Section 8.2 of Gill, Murray, and Wright [1981].

The following abbreviations are used in the tables to describe the conditions under which the algorithm terminates : †

OPT.	-	optimal point found
*	-	current point cannot be improved
**	-	optimal solution found, but requested accuracy could not be achieved
F LIM.	-	function evaluation limit reached

† A '*' corresponds to the situation in which the algorithm terminates due to failure in the linesearch to find an acceptable step at the current iteration. A '**' occurs when condition (5.3.1) is satisfied but not condition (5.3.2);that is, conditions for optimality are met at the current point but the iterates have not yet converged.

Numerical Results for NPSOL

	n	m	Opt Tol	Step Lim	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
1.^{0}	2	2	10^{-10}		27	22	1.41	10^{-6}	10^{-6}	10^{-13}	OPT.
			10^{-14}		29	24	1.41	10^{-11}	10^{-9}	10^{-21}	OPT.
2.^{0}	2	2	10^{-10}		9	6	6.40	10^{-8}	10^{-7}	10^{-15}	OPT.
			10^{-14}		10	7	6.40	10^{-10}	10^{-9}	10^{-20}	OPT.
3.^{0}	2	2	10^{-10}		897	283	9.11	10^{-9}	10^{-7}	10^{-18}	OPT.
			10^{-14}		897	283	9.11	10^{-9}	10^{-7}	10^{-18}	OPT.
4.^{0}	2	3	10^{-10}		20	14	10^{6}	10^{-9}	10^{-3}	10^{-18}	*
			10^{-14}		20	14	10^{6}	10^{-9}	10^{-3}	10^{-18}	*
5.^{0}	2	3	10^{-10}		20	15	3.04	10^{-6}	10^{-6}	10^{-13}	OPT.
			10^{-14}		22	17	3.04	10^{-10}	10^{-10}	10^{-20}	OPT.
6.	2	10	10^{-10}	0.1	14	9	.365	10^{1}	10^{-5}	10^{-6}	OPT.
			10^{-14}	0.1	15	10	.365	10^{1}	10^{-8}	10^{-6}	OPT.
7.^{0}	3	3	10^{-10}		37	31	1.00	10^{-8}	10^{-6}	10^{-15}	OPT.
			10^{-14}		38	32	1.00	10^{-11}	10^{-10}	10^{-21}	OPT.
8.	3	15	10^{-10}		22	16	2.60	10^{-1}	10^{-8}	10^{-8}	OPT.
			10^{-14}		23	17	2.60	10^{-1}	10^{-10}	10^{-8}	OPT.
9.	3	15	10^{-10}		8	4	1.08	10^{-4}	10^{-8}	10^{-14}	OPT.
			10^{-14}		9	5	1.08	10^{-4}	10^{-12}	10^{-14}	OPT.
10.	3	16	10^{-10}		450	328	10^{3}	10^{1}	10^{-1}	10^{-6}	*
			10^{-14}		450	328	10^{3}	10^{1}	10^{-1}	10^{-6}	*
11.^{0}	3	10	10^{-10}		2	1	8.39	10^{-1}	0.00	10^{-2}	OPT.
			10^{-14}		2	1	8.39	10^{-1}	0.00	10^{-2}	OPT.
12.^{0}	3	10	10^{-10}		34	27	10.1	10^{-8}	10^{-7}	10^{-15}	OPT.
			10^{-14}		35	28	10.1	10^{-9}	10^{-9}	10^{-18}	OPT.
13.^{0}	4	4	10^{-10}		66	59	10^{-5}	10^{-9}	10^{-10}	10^{-18}	OPT.
			10^{-14}		71	64	10^{-5}	10^{-9}	10^{-12}	10^{-18}	OPT.
14.^{0}	4	6	10^{-10}		50	39	2.00	10^{-8}	10^{-6}	10^{-15}	OPT.
			10^{-14}		51	40	2.00	10^{-9}	10^{-8}	10^{-18}	OPT.
15.	4	11	10^{-10}		33	20	.328	10^{-2}	10^{-8}	10^{-9}	OPT.
			10^{-14}		35	22	.328	10^{-2}	10^{-10}	10^{-9}	OPT.
16.	4	20	10^{-10}		24	17	17.6	10^{2}	10^{-3}	10^{-8}	OPT.
			10^{-14}		25	18	17.6	10^{2}	10^{-4}	10^{-8}	OPT.
17.	5	33	10^{-10}	0.1	32	20	2.37	10^{-2}	10^{-5}	10^{-7}	OPT.
			10^{-14}	0.1	56	44	2.46	10^{-2}	10^{-9}	10^{-11}	OPT.
18.^{0}	6	13	10^{-10}		45	39	18.7	10^{-1}	10^{-8}	10^{-2}	OPT.
			10^{-14}		45	39	18.7	10^{-1}	10^{-8}	10^{-2}	OPT.
19.	11	65	10^{-10}		88	59	9.38	10^{-1}	10^{-5}	10^{-8}	OPT.
			10^{-14}		90	61	9.38	10^{-1}	10^{-7}	10^{-8}	OPT.

Numerical Results for NPSOL

	n	m	Opt Tol	Step Lim	f, J evals.	iters.	$\lVert x^*\rVert_2$	$\lVert f^*\rVert_2$	$\lVert g^*\rVert_2$	est. err.	conv.
20a.	6	31	10^{-10}		43	36	2.44	10^{-2}	10^{-8}	10^{-10}	OPT.
			10^{-14}		46	39	2.44	10^{-2}	10^{-10}	10^{-10}	OPT.
20b.	9	31	10^{-10}		83	75	6.06	10^{-3}	10^{-8}	10^{-13}	OPT.
			10^{-14}		85	77	6.06	10^{-3}	10^{-10}	10^{-13}	OPT.
20c.	12	31	10^{-10}		55	48	1.06	10^{-4}	10^{-6}	10^{-7}	OPT.
			10^{-14}		151	129	16.6	10^{-5}	10^{-11}	10^{-16}	OPT.
20d.	20	31	10^{-10}		73	63	1.06	10^{-4}	10^{-7}	10^{-8}	OPT.
			10^{-14}		114	93	1.06	10^{-6}	10^{-9}	10^{-12}	OPT.
21a.^{0}	10	10	10^{-10}		101	59	3.16	10^{-6}	10^{-5}	10^{-12}	OPT.
			10^{-14}		104	62	3.16	10^{-9}	10^{-8}	10^{-17}	OPT.
21b.^{0}	20	20	10^{-10}		252	157	4.47	10^{-5}	10^{-5}	10^{-11}	OPT.
			10^{-14}		265	170	4.47	10^{-9}	10^{-8}	10^{-17}	OPT.
22a.^{0}	12	12	10^{-10}		83	72	10^{-3}	10^{-6}	10^{-8}	10^{-11}	OPT.
			10^{-14}		165	154	10^{-4}	10^{-8}	10^{-12}	10^{-17}	OPT.
22b.^{0}	20	20	10^{-10}		103	85	10^{-3}	10^{-6}	10^{-7}	10^{-11}	OPT.
			10^{-14}		196	178	10^{-4}	10^{-7}	10^{-11}	10^{-14}	OPT.
23a.	4	5	10^{-10}		198	142	.500	10^{-3}	10^{-10}	10^{-10}	OPT.
			10^{-14}		198	142	.500	10^{-3}	10^{-10}	10^{-10}	OPT.
23b.	10	11	10^{-10}		117	79	.500	10^{-2}	10^{-9}	10^{-11}	OPT.
			10^{-14}		124	86	.500	10^{-2}	10^{-12}	10^{-11}	OPT.
24a.	4	8	10^{-10}		23	15	.828	10^{-3}	10^{-8}	10^{-6}	OPT.
			10^{-14}		462	346	.759	10^{-3}	10^{-11}	10^{-11}	OPT.
24b.	10	20	10^{-10}		368	263	.598	10^{-2}	10^{-6}	10^{-9}	OPT.
			10^{-14}		419	314	.598	10^{-2}	10^{-9}	10^{-9}	OPT.
25a.^{0}	10	12	10^{-10}		19	17	3.16	10^{-7}	10^{-8}	10^{-14}	OPT.
			10^{-14}		20	18	3.16	10^{-14}	10^{-13}	10^{-28}	OPT.
25b.^{0}	20	22	10^{-10}		24	22	4.47	10^{-13}	10^{-11}	10^{-25}	OPT.
			10^{-14}		24	22	4.47	10^{-13}	10^{-11}	10^{-25}	OPT.
26a.^{0}	10	10	10^{-10}		33	31	.328	10^{-2}	10^{-8}	10^{-5}	OPT.
			10^{-14}		35	33	.328	10^{-2}	10^{-9}	10^{-5}	OPT.
26b.^{0}	20	20	10^{-10}		76	44	.231	10^{-3}	10^{-8}	10^{-5}	OPT.
			10^{-14}		78	46	.231	10^{-3}	10^{-9}	10^{-5}	OPT.
27a.^{0}	10	10	10^{-10}		18	15	3.16	10^{-9}	10^{-8}	10^{-18}	OPT.
			10^{-14}		19	16	3.16	10^{-10}	10^{-9}	10^{-21}	OPT.
27b.^{0}	20	20	10^{-10}		30	17	4.47	10^{-7}	10^{-6}	10^{-13}	OPT.
			10^{-14}		32	19	4.47	10^{-9}	10^{-9}	10^{-18}	OPT.
28a.^{0}	10	10	10^{-10}		33	18	.412	10^{-7}	10^{-7}	10^{-13}	OPT.
			10^{-14}		36	20	.412	10^{-8}	10^{-8}	10^{-15}	OPT.
28b.^{0}	20	20	10^{-10}		54	40	.571	10^{-8}	10^{-9}	10^{-16}	OPT.
			10^{-14}		56	42	.571	10^{-11}	10^{-10}	10^{-21}	OPT.

	n	m	Opt Tol	Step Lim	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
29a.[0]	10	10	10^{-10}		7	6	.412	10^{-8}	10^{-8}	10^{-15}	OPT.
			10^{-14}		8	7	.412	10^{-9}	10^{-9}	10^{-18}	OPT.
29b.[0]	20	20	10^{-10}		7	6	.571	10^{-7}	10^{-7}	10^{-15}	OPT.
			10^{-14}		8	7	.571	10^{-9}	10^{-9}	10^{-18}	OPT.
30a.[0]	10	10	10^{-10}		37	19	2.05	10^{-6}	10^{-5}	10^{-12}	OPT.
			10^{-14}		40	22	2.05	10^{-9}	10^{-8}	10^{-17}	OPT.
30b.[0]	20	20	10^{-10}		65	33	3.22	10^{0}	10^{-5}	10^{0}	OPT.
			10^{-14}		67	35	3.22	10^{0}	10^{-7}	10^{0}	OPT.
31a.[0]	10	10	10^{-10}		67	38	2.18	10^{0}	10^{-5}	10^{0}	OPT.
			10^{-14}		70	41	2.18	10^{0}	10^{-7}	10^{0}	OPT.
31b.[0]	20	20	10^{-10}		141	88	1.90	10^{0}	10^{-5}	10^{0}	OPT.
			10^{-14}		144	91	1.90	10^{0}	10^{-7}	10^{0}	OPT.
32.[L]	10	20	10^{-10}		2	1	3.16	10^{0}	10^{-15}	0.00	OPT.
			10^{-14}		2	1	3.16	10^{0}	10^{-15}	0.00	OPT.
33.[L]	10	20	10^{-10}		4	2	1.46	10^{0}	10^{-10}	10^{-6}	**
			10^{-14}		4	2	1.46	10^{0}	10^{-10}	10^{-6}	**
34.[L]	10	20	10^{-10}		4	2	1.78	10^{0}	10^{-11}	10^{-6}	**
			10^{-14}		4	2	1.78	10^{0}	10^{-11}	10^{-6}	**
35a.	8	8	10^{-10}		31	20	1.65	10^{-1}	10^{-6}	10^{-9}	OPT.
			10^{-14}		33	22	1.65	10^{-1}	10^{-8}	10^{-9}	OPT.
35b.[0]	9	9	10^{-10}		29	16	1.73	10^{-5}	10^{-5}	10^{-10}	OPT.
			10^{-14}		32	19	1.73	10^{-7}	10^{-7}	10^{-13}	OPT.
35c.	10	10	10^{-10}		37	26	1.81	10^{-1}	10^{-6}	10^{-8}	OPT.
			10^{-14}		42	31	1.81	10^{-1}	10^{-9}	10^{-8}	OPT.
36a.[0]	4	4	10^{-10}		609	456	8.87	10^{-5}	10^{-6}	10^{-9}	OPT.
			10^{-14}		3762	2712	17.1	10^{-6}	10^{-6}	10^{-11}	*
36b.[0]	9	9	10^{-10}		887	660	10.1	10^{-5}	10^{-6}	10^{-9}	OPT.
			10^{-14}		3211	2357	16.3	10^{-6}	10^{-6}	10^{-11}	*
36c.[0]	9	9	10^{-10}		3	1	1.73	0.00	0.00	0.00	OPT.
			10^{-14}		3	1	1.73	0.00	0.00	0.00	OPT.
36d.[0]	9	9	10^{-10}		1063	813	10.7	10^{-5}	10^{-6}	10^{-10}	OPT.
			10^{-14}		3680	2734	16.9	10^{-6}	10^{-6}	10^{-11}	*
37.	2	16	10^{-10}		14	8	10.8	10^{1}	10^{-5}	10^{-6}	OPT.
			10^{-14}		15	9	10.8	10^{1}	10^{-8}	10^{-6}	OPT.
38.	3	16	10^{-10}	0.01	21	14	26.1	10^{1}	10^{-5}	10^{-6}	OPT.
			10^{-14}	0.01	23	16	26.1	10^{1}	10^{-8}	10^{-6}	OPT.

Numerical Results for NPSOL

	n	m	Opt Tol	Step Lim	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
39a.	2	3	10^{-10}		10	9	10^{-6}	10^{-1}	10^{-8}	10^{-7}	OPT.
			10^{-14}		11	10	10^{-6}	10^{-1}	10^{-11}	10^{-7}	OPT.
39b.	2	3	10^{-10}		9	8	10^{-7}	10^{-1}	10^{-7}	10^{-7}	OPT.
			10^{-14}		10	9	10^{-7}	10^{-1}	10^{-9}	10^{-7}	OPT.
39c.	2	3	10^{-10}		6	5	10^{-7}	10^{-1}	10^{-7}	10^{-7}	OPT.
			10^{-14}		8	7	10^{-7}	10^{-1}	10^{-16}	10^{-7}	OPT.
39d.	2	3	10^{-10}		10	7	10^{-7}	10^{-1}	10^{-7}	10^{-7}	OPT.
			10^{-14}		11	8	10^{-7}	10^{-1}	10^{-10}	10^{-7}	OPT.
39e.	2	3	10^{-10}		16	11	10^{-7}	10^{-1}	10^{-6}	10^{-7}	OPT.
			10^{-14}		17	12	10^{-8}	10^{-1}	10^{-8}	10^{-7}	OPT.
39f.	2	3	10^{-10}		28	21	10^{-9}	10^{-1}	10^{-6}	10^{-7}	OPT.
			10^{-14}		29	22	10^{-9}	10^{-1}	10^{-8}	10^{-7}	OPT.
39g.	2	3	10^{-10}		30	21	10^{-10}	10^{-1}	10^{-8}	10^{-7}	OPT.
			10^{-14}		31	22	10^{-10}	10^{-1}	10^{-13}	10^{-7}	OPT.
40a.	3	4	10^{-10}		11	10	10^{-6}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-14}		12	11	10^{-6}	10^{0}	10^{-10}	10^{-7}	OPT.
40b.	3	4	10^{-10}		11	10	10^{-6}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-14}		12	11	10^{-6}	10^{0}	10^{-11}	10^{-7}	OPT.
40c.	3	4	10^{-10}		9	7	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-14}		10	8	10^{-7}	10^{0}	10^{-11}	10^{-7}	OPT.
40d.	3	4	10^{-10}		14	9	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-14}		15	10	10^{-7}	10^{0}	10^{-11}	10^{-7}	OPT.
40e.	3	4	10^{-10}		19	11	10^{-6}	10^{0}	10^{-6}	10^{-7}	OPT.
			10^{-14}		20	12	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
40f.	3	4	10^{-10}		33	23	10^{-8}	10^{0}	10^{-6}	10^{-7}	OPT.
			10^{-14}		34	24	10^{-8}	10^{0}	10^{-8}	10^{-7}	OPT.
40g.	3	4	10^{-10}		45	35	10^{-9}	10^{0}	10^{-5}	10^{-7}	OPT.
			10^{-14}		46	36	10^{-9}	10^{0}	10^{-7}	10^{-7}	OPT.
41a.	5	10	10^{-10}		12	7	10^{-6}	10^{0}	10^{-11}	10^{-7}	OPT.
			10^{-14}		12	7	10^{-6}	10^{0}	10^{-11}	10^{-7}	OPT.
41b.	5	10	10^{-10}		12	7	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
			10^{-14}		13	8	10^{-6}	10^{0}	10^{-12}	10^{-7}	OPT.
41c.	5	10	10^{-10}		12	7	10^{-6}	10^{0}	10^{-7}	10^{-7}	OPT.
			10^{-14}		14	9	10^{-6}	10^{0}	10^{-10}	10^{-7}	OPT.
41d.	5	10	10^{-10}		17	12	10^{-6}	10^{0}	10^{-6}	10^{-7}	OPT.
			10^{-14}		20	15	10^{-6}	10^{0}	10^{-11}	10^{-7}	OPT.
41e.	5	10	10^{-10}		51	35	10^{-6}	10^{0}	10^{-5}	10^{-7}	OPT.
			10^{-14}		54	38	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
41f.	5	10	10^{-10}		51	30	10^{-7}	10^{0}	10^{-5}	10^{-7}	OPT.
			10^{-14}		53	32	10^{-7}	10^{0}	10^{-7}	10^{-7}	OPT.
41g.	5	10	10^{-10}		62	45	10^{-8}	10^{0}	10^{-5}	10^{-7}	OPT.
			10^{-14}		69	52	10^{-8}	10^{0}	10^{-7}	10^{-7}	OPT.

	n	m	Opt Tol	Step Lim	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
42a.^{0}	4	24	10^{-10}	1.0	60	43	61.9	10^{-8}	10^{-6}	10^{-17}	OPT
			10^{-14}	1.0	61	44	61.9	10^{-10}	10^{-8}	10^{-20}	OPT.
42b.^{0}	4	24	10^{-10}	1.0	140	85	81.4	10^{-8}	10^{-6}	10^{-16}	OPT
			10^{-14}	1.0	141	86	81.4	10^{-11}	10^{-9}	10^{-23}	OPT.
42c.^{0}	4	24	10^{-10}	0.1	51	49	60.3	10^{-8}	10^{-5}	10^{-16}	OPT.
			10^{-14}	0.1	52	50	60.3	10^{-10}	10^{-7}	10^{-20}	OPT.
42d.^{0}	4	24	10^{-10}	0.1	56	51	60.3	10^{-9}	10^{-7}	10^{-19}	OPT.
			10^{-14}	0.1	57	52	60.3	10^{-11}	10^{-9}	10^{-23}	OPT.
43a.^{0}	5	16	10^{-10}	0.01	44	38	53.6	10^{-1}	10^{-6}	10^{-2}	OPT.
			10^{-14}	0.01	53	47	53.6	10^{-1}	10^{-9}	10^{-2}	OPT.
43b.^{0}	5	16	10^{-10}	1.0	37	25	46.2	10^{1}	10^{-4}	10^{2}	OPT.
			10^{-14}	1.0	38	26	46.2	10^{1}	10^{-5}	10^{2}	OPT
43c.^{0}	5	16	10^{-10}	0.01	44	39	53.6	10^{-1}	10^{-6}	10^{-2}	OPT.
			10^{-14}	0.01	54	49	53.6	10^{-1}	10^{-9}	10^{-2}	OPT.
43d.^{0}	5	16	10^{-10}	0.01	112	81	53.6	10^{-1}	10^{-5}	10^{-2}	OPT.
			10^{-14}	0.01	120	89	53.6	10^{-1}	10^{-8}	10^{-2}	OPT.
43e.^{0}	5	16	10^{-10}	0.01	95	62	54.0	10^{-8}	10^{-6}	10^{-15}	OPT.
			10^{-14}	0.01	97	64	54.0	10^{-10}	10^{-8}	10^{-20}	OPT.
43f.^{0}	5	16	10^{-10}	1.0	56	43	54.0	10^{-8}	10^{-5}	10^{-15}	OPT.
			10^{-14}	1.0	59	46	54.0	10^{-11}	10^{-9}	10^{-21}	OPT.
44a.^{0}	6	6	10^{-10}		488	383	4.03	10^{-8}	10^{-6}	10^{-15}	OPT.
			10^{-14}		490	385	4.03	10^{-9}	10^{-8}	10^{-18}	OPT.
44b.^{0}	6	6	10^{-10}		57	35	3.52	10^{-7}	10^{-5}	10^{-13}	OPT.
			10^{-14}		59	37	3.52	10^{-10}	10^{-9}	10^{-19}	OPT.
44c.^{0}	6	6	10^{-10}		(6001)	(4430)	10^{3}	10^{0}	10^{-1}	10^{0}	F LIM.
			10^{-14}		(6001)	(4430)	10^{3}	10^{0}	10^{-1}	10^{0}	F LIM.
44d.^{0}	6	6	10^{-10}		(6000)	(3944)	13.9	10^{-2}	10^{0}	10^{-3}	F LIM.
			10^{-14}		(6000)	(3944)	13.9	10^{-2}	10^{0}	10^{-3}	F LIM.
44e.^{0}	6	6	10^{-10}		1976	1420	9.27	10^{-8}	10^{-6}	10^{-17}	OPT.
			10^{-14}		1978	1422	9.27	10^{-11}	10^{-9}	10^{-22}	OPT.
45a.^{0}	8	8	10^{-10}		474	371	4.06	10^{-7}	10^{-7}	10^{-14}	OPT.
			10^{-14}		476	373	4.06	10^{-9}	10^{-8}	10^{-19}	OPT.
45b.^{0}	8	8	10^{-10}		82	57	3.56	10^{-7}	10^{-6}	10^{-13}	OPT.
			10^{-14}		84	59	3.56	10^{-9}	10^{-8}	10^{-17}	OPT.
45c.^{0}	8	8	10^{-10}		(8000)	(6180)	12.8	10^{-1}	10^{1}	10^{-1}	F LIM.
			10^{-14}		(8000)	(6180)	12.8	10^{-1}	10^{1}	10^{-1}	F LIM.
45d.^{0}	8	8	10^{-10}		1654	1302	15.3	10^{-9}	10^{-5}	10^{-17}	OPT.
			10^{-14}		1656	1304	15.3	10^{-11}	10^{-7}	10^{-21}	OPT.
45e.^{0}	8	8	10^{-10}		(8000)	(5432)	67.8	10^{-1}	10^{2}	10^{-2}	F LIM.
			10^{-14}		(8000)	(5432)	67.8	10^{-1}	10^{2}	10^{-2}	F LIM.

5.4 Second-Derivative (Modified-Newton) Linesearch Method (NPL/NAG MNA)

5.4.1 Software and Algorithm

The results were obtained using subroutine MNA from the National Physical Laboratory, available at Stanford Linear Accelerator Center. The algorithm implements a modified Newton method in which the search direction at each iteration is the solution to a subproblem of the form

$$\min_{p \in \Re^n} g_k^T p + \frac{1}{2} p^T H_k p,$$

and the exact Hessian matrix is replaced by modified Cholesky factors if it is either indefinite or computationally singular (see Gill and Murray [1974a] and Section 2.4.1). A step length along the search direction is then computed by a linesearch method [Gill and Murray (1974b)] that uses both function and gradient information to obtain sufficient decrease in the objective function. MNA requires exact second derivatives, and is similar to subroutine E04LBF from the NAG Library [1984], the principal difference being that the latter allows specification of fixed upper and lower bounds on the variables.

5.4.2 Parameters

Parameters were kept at their default values with the following exceptions :

MAXCAL	-	min $\{9999, 1000n\}$	function evaluation limit
XTOL	-	varied; see tables	accuracy in x
ETA	-	0.9	linesearch accuracy
STEPMX	-	usually 10^6 (default) †	maximum step for linesearch

† In some cases the default STEPMX $= 10^6$ was too large and overflow occurred during function evaluation in the linesearch. These cases are indicated in the table by giving the lower value of STEPMX that was subsequently used to obtain the results in the column labeled "max. step".

See NAG [1984] for details concerning the parameters.

5.4.3 Convergence Criteria

The following quantities will be used in describing the convergence criteria:

objective function	:	$\mathcal{F}_k \ \ (= \frac{1}{2} f_k^T f_k)$
objective gradient	:	$g_k = \nabla \mathcal{F}_k \ \ (= J_k^T f_k)$
search direction	:	p_k, the minimizer of the subproblem
steplength	:	α_k, determined by the linesearch

An iterate is determined to be optimal by **MNA** if the following four conditions hold:

$$\alpha_k \|p_k\|_2 (\text{XTOL} + \sqrt{\epsilon_M})(1 + \|x_k\|_2) \tag{5.4.1}$$

and

$$\mathcal{F}_{k-1} - \mathcal{F}_k < (\text{XTOL}^2 + \epsilon_M)(1 + |\mathcal{F}_k|) \tag{5.4.2}$$

and

$$\|g_k\|_2 < (\text{XTOL} + \epsilon_M^{1/3})(1 + |\mathcal{F}_k|) \tag{5.4.3}$$

and

$$\nabla^2 \mathcal{F}_k \ \text{ is positive definite,} \tag{5.4.4}$$

or if

$$\|g_k\|_2 < 0.01\sqrt{\epsilon_M}. \tag{5.4.5}$$

A necessary condition for optimality is that the gradient vanish, and conditions (5.4.3) and (5.4.5) are intended to test whether this requirement is approximately satisfied at x_k. Conditions (5.4.1) and (5.4.2) are meant to ensure that the sequence $\{x_k\}$ has converged, while condition (5.4.4), together with condition (5.4.3), implies that sufficient conditions for a strict local minimum appear to hold at x_k. Condition (5.4.5) allows **MNA** to accept a point as a local mimimum if a more restrictive test than (5.4.1) on the necessary condition is met, but one or more of the other conditions for convergence do not hold. For a detailed discussion of convergence criteria similar to these, see Section 8.2 of Gill, Murray, and Wright [1981].

The following abbreviations are used in the tables to describe the conditions under which the algorithm terminates:

OPT.	-	optimal point found
*	-	current point cannot be improved †
F LIM.	-	function evaluation limit reached
TIME	-	time limit exceeded

† A '*' corresponds to the situation in which the algorithm terminates due to failure in the linesearch to find an acceptable step at the current iteration.

Numerical Results for MNA

	n	m	XTOL	max. step	f, J evals.	iters.	$\|\|x^*\|\|_2$	$\|\|f^*\|\|_2$	$\|\|g^*\|\|_2$	est. err.	conv.
1.^{0}	2	2	10^{-6}		14	10	1.41	10^{-15}	10^{-15}	10^{-30}	OPT
			10^{-10}		14	10	1.41	10^{-15}	10^{-15}	10^{-30}	OPT.
2.^{0}	2	2	10^{-6}		8	6	6.40	10^{-9}	10^{-8}	10^{-18}	OPT
			10^{-10}		8	6	6.40	10^{-9}	10^{-8}	10^{-18}	OPT.
3.^{0}	2	2	10^{-6}		175	65	9.11	10^{-9}	10^{-4}	10^{-19}	*
			10^{-10}		175	65	9.11	10^{-9}	10^{-4}	10^{-19}	*
4.^{0}	2	3	10^{-6}		1	1	10^{6}	10^{-16}	10^{-10}	10^{-32}	*
			10^{-10}		1	1	10^{6}	10^{-16}	10^{-10}	10^{-32}	*
5.^{0}	2	3	10^{-6}	10^{3}	35	9	3.04	10^{-10}	10^{-9}	10^{-19}	*
			10^{-10}	10^{3}	35	9	3.04	10^{-10}	10^{-9}	10^{-19}	*
6.	2	10	10^{-6}		12	9	.365	10^{1}	10^{-7}	10^{-6}	*
			10^{-10}		12	9	.365	10^{1}	10^{-7}	10^{-6}	*
7.^{0}	3	3	10^{-6}		14	11	1.00	10^{-12}	10^{-12}	10^{-25}	OPT.
			10^{-10}		14	11	1.00	10^{-12}	10^{-12}	10^{-25}	OPT.
8.	3	15	10^{-6}		11	10	2.60	10^{-1}	10^{-15}	10^{-8}	OPT.
			10^{-10}		11	10	2.60	10^{-1}	10^{-15}	10^{-8}	OPT.
9.	3	15	10^{-6}		3	2	1.08	10^{-4}	10^{-11}	10^{-14}	OPT.
			10^{-10}		3	2	1.08	10^{-4}	10^{-11}	10^{-14}	OPT.
10.	3	16	10^{-6}		249	164	10^{4}	10^{1}	10^{-1}	10^{-6}	*
			10^{-10}		249	164	10^{4}	10^{1}	10^{-1}	10^{-6}	*
11.^{0}	3	10	10^{-6}		538	341	55.9	10^{-11}	10^{-11}	10^{-21}	OPT.
			10^{-10}		538	341	55.9	10^{-11}	10^{-11}	10^{-21}	OPT.
12.^{0}	3	10	10^{-6}		43	18	10.1	10^{-16}	10^{-16}	10^{-31}	OPT.
			10^{-10}		43	18	10.1	10^{-16}	10^{-16}	10^{-31}	OPT.
13.^{0}	4	4	10^{-6}		23	22	10^{-4}	10^{-7}	10^{-10}	10^{-15}	OPT.
			10^{-10}		23	22	10^{-4}	10^{-7}	10^{-10}	10^{-15}	OPT.
14.^{0}	4	6	10^{-6}		54	26	2.00	10^{-13}	10^{-12}	10^{-26}	OPT.
			10^{-10}		54	26	2.00	10^{-13}	10^{-12}	10^{-26}	OPT.
15.	4	11	10^{-6}		20	7	.328	10^{-2}	10^{-13}	10^{-9}	OPT
			10^{-10}		20	7	.328	10^{-2}	10^{-13}	10^{-9}	OPT.
16.	4	20	10^{-6}		9	8	17.6	10^{2}	10^{-10}	10^{-8}	OPT.
			10^{-10}		10	9	17.6	10^{2}	10^{-10}	10^{-8}	*
17.	5	33	10^{-6}		43	28	2.46	10^{-2}	10^{-14}	10^{-11}	OPT.
			10^{-10}		43	28	2.46	10^{-2}	10^{-14}	10^{-11}	*
18.^{0}	6	13	10^{-6}		44	20	12.3	10^{-12}	10^{-13}	10^{-23}	OPT.
			10^{-10}		44	20	12.3	10^{-12}	10^{-13}	10^{-23}	OPT.
19.	11	65	10^{-6}	10.0	7	3	9.38	10^{-1}	10^{-9}	10^{-8}	OPT.
			10^{-10}	10.0	8	4	9.38	10^{-1}	10^{-11}	10^{-8}	OPT.

Numerical Results for MNA

	n	m	XTOL	max. step	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
20a.	6	31	10^{-6}		13	12	2.44	10^{-2}	10^{-12}	10^{-10}	OPT.
			10^{-10}		13	12	2.44	10^{-2}	10^{-12}	10^{-10}	OPT.
20b.	9	31	10^{-6}		14	13	6.06	10^{-3}	10^{-14}	10^{-13}	OPT.
			10^{-10}		14	13	6.06	10^{-3}	10^{-14}	10^{-13}	OPT.
20c.	12	31	10^{-6}		14	13	16.6	10^{-5}	10^{-13}	10^{-16}	OPT.
			10^{-10}		14	13	16.6	10^{-5}	10^{-13}	10^{-16}	OPT.
20d.	20	31	10^{-6}		1295	679	10^{6}	10^{-3}	10^{-7}	10^{-5}	*
			10^{-10}		1295	679	10^{6}	10^{-3}	10^{-7}	10^{-5}	*
21a.[0]	10	10	10^{-6}		14	10	3.16	10^{-15}	10^{-15}	10^{-29}	OPT.
			10^{-10}		14	10	3.16	10^{-15}	10^{-15}	10^{-29}	OPT.
21b.[0]	20	20	10^{-6}		14	10	4.47	10^{-14}	10^{-14}	10^{-29}	OPT.
			10^{-10}		14	10	4.47	10^{-14}	10^{-14}	10^{-29}	OPT.
22a.[0]	12	12	10^{-6}		23	22	10^{-4}	10^{-7}	10^{-10}	10^{-14}	OPT.
			10^{-10}		23	22	10^{-4}	10^{-7}	10^{-10}	10^{-14}	OPT.
22b.[0]	20	20	10^{-6}		24	23	10^{-4}	10^{-7}	10^{-11}	10^{-15}	OPT.
			10^{-10}		24	23	10^{-4}	10^{-7}	10^{-11}	10^{-15}	OPT.
23a.	4	5	10^{-6}		43	34	.500	10^{-3}	10^{-10}	10^{-10}	OPT.
			10^{-10}		43	34	.500	10^{-3}	10^{-10}	10^{-10}	OPT.
23b.	10	11	10^{-6}		44	36	.500	10^{-2}	10^{-12}	10^{-11}	OPT.
			10^{-10}		44	36	.500	10^{-2}	10^{-12}	10^{-11}	OPT.
24a.	4	8	10^{-6}		158	110	.759	10^{-3}	10^{-10}	10^{-11}	OPT.
			10^{-10}		158	110	.759	10^{-3}	10^{-10}	10^{-11}	OPT.
24b.	10	20	10^{-6}		133	94	.598	10^{-2}	10^{-13}	10^{-9}	OPT.
			10^{-10}		133	94	.598	10^{-2}	10^{-13}	10^{-9}	OPT.
25a.[0]	10	12	10^{-6}		14	13	3.16	10^{-10}	10^{-8}	10^{-19}	*
			10^{-10}		14	13	3.16	10^{-10}	10^{-8}	10^{-19}	*
25b.[0]	20	22	10^{-6}		17	16	4.47	10^{-8}	10^{-6}	10^{-15}	OPT.
			10^{-10}		18	17	4.47	10^{-8}	10^{-6}	10^{-15}	*
26a.[0]	10	10	10^{-6}		21	11	.306	10^{-15}	10^{-16}	10^{-30}	OPT.
			10^{-10}		21	11	.306	10^{-15}	10^{-16}	10^{-30}	OPT.
26b.[0]	20	20	10^{-6}		30	13	.189	10^{-13}	10^{-13}	10^{-25}	OPT.
			10^{-10}		30	13	.189	10^{-13}	10^{-13}	10^{-25}	OPT.
27a.[0]	10	10	10^{-6}		22	12	3.16	10^{-12}	10^{-12}	10^{-24}	OPT.
			10^{-10}		22	12	3.16	10^{-12}	10^{-12}	10^{-24}	OPT.
27b.[0]	20	20	10^{-6}		31	15	4.47	10^{-13}	10^{-13}	10^{-27}	OPT.
			10^{-10}		31	15	4.47	10^{-13}	10^{-13}	10^{-27}	OPT.
28a.[0]	10	10	10^{-6}		4	3	.412	10^{-12}	10^{-13}	10^{-24}	OPT.
			10^{-10}		4	3	.412	10^{-12}	10^{-13}	10^{-24}	OPT.
28b.[0]	20	20	10^{-6}		4	3	.571	10^{-13}	10^{-14}	10^{-25}	OPT.
			10^{-10}		4	3	.571	10^{-13}	10^{-14}	10^{-25}	OPT.

	n	m	XTOL	max. step	f, J evals.	iters.	$\Vert x^*\Vert_2$	$\Vert f^*\Vert_2$	$\Vert g^*\Vert_2$	est. err.	conv.
29a.^{0}	10	10	10^{-6}		4	3	.412	10^{-10}	10^{-10}	10^{-20}	OPT.
			10^{-10}		4	3	.412	10^{-10}	10^{-10}	10^{-20}	OPT.
29b.^{0}	20	20	10^{-6}		4	3	.571	10^{-10}	10^{-10}	10^{-20}	OPT.
			10^{-10}		4	3	.571	10^{-10}	10^{-10}	10^{-20}	OPT.
30a.^{0}	10	10	10^{-6}		7	6	2.05	10^{-16}	10^{-15}	10^{-31}	OPT.
			10^{-10}		7	6	2.05	10^{-16}	10^{-15}	10^{-31}	OPT.
30b.^{0}	20	20	10^{-6}		7	6	3.04	10^{-15}	10^{-14}	10^{-30}	OPT.
			10^{-10}		7	6	3.04	10^{-15}	10^{-14}	10^{-30}	OPT.
31a.^{0}	10	10	10^{-6}		9	8	1.80	10^{-13}	10^{-12}	10^{-26}	OPT.
			10^{-10}		9	8	1.80	10^{-13}	10^{-12}	10^{-26}	OPT.
31b.^{0}	20	20	10^{-6}		9	8	2.66	10^{-13}	10^{-12}	10^{-26}	OPT.
			10^{-10}		9	8	2.66	10^{-13}	10^{-12}	10^{-26}	OPT.
32.^{L}	10	20	10^{-6}		4	2	3.16	10^{0}	10^{-15}	10^{-15}	OPT.
			10^{-10}		4	2	3.16	10^{0}	10^{-15}	10^{-15}	OPT.
33.^{L}	10	20	10^{-6}		27	1	10^{4}	10^{0}	10^{-8}	10^{-6}	*
			10^{-10}		27	1	10^{4}	10^{0}	10^{-8}	10^{-6}	*
34.^{L}	10	20	10^{-6}		20	1	21.2	10^{0}	10^{-10}	10^{-6}	*
			10^{-10}		20	1	21.2	10^{0}	10^{-10}	10^{-6}	*
35a.	8	8	10^{-6}		41	15	1.65	10^{-1}	10^{-13}	10^{-9}	OPT.
			10^{-10}		41	15	1.65	10^{-1}	10^{-13}	10^{-9}	OPT.
35b.^{0}	9	9	10^{-6}		66	19	1.73	10^{-10}	10^{-9}	10^{-20}	*
			10^{-10}		66	19	1.73	10^{-10}	10^{-9}	10^{-20}	*
35c.	10	10	10^{-6}		86	17	1.76	10^{-1}	10^{-9}	10^{-9}	*
			10^{-10}		86	17	1.76	10^{-1}	10^{-9}	10^{-9}	*
36a.^{0}	4	4	10^{-6}		(4002)	(2658)	51.4	10^{-9}	10^{-9}	10^{-18}	F LIM.
			10^{-10}		(4002)	(2658)	51.4	10^{-9}	10^{-9}	10^{-18}	F LIM.
36b.^{0}	9	9	10^{-6}		(9014)	(692)	51.2	10^{-7}	10^{-5}	10^{-13}	F LIM.
			10^{-10}		(9014)	(692)	51.2	10^{-7}	10^{-5}	10^{-13}	F LIM.
36c.^{0}	9	9	10^{-6}		3188	310	1.73	10^{-11}	10^{-11}	10^{-22}	OPT.
			10^{-10}		3188	310	1.73	10^{-11}	10^{-11}	10^{-22}	OPT.
36d.^{0}	9	9	10^{-6}		(9003)	(1985)	990.	10^{-3}	10^{-2}	10^{-6}	F LIM.
			10^{-10}		(9003)	(1985)	990.	10^{-3}	10^{-2}	10^{-6}	F LIM.
37.	2	16	10^{-6}		6	5	8.85	10^{1}	10^{-10}	10^{-6}	OPT.
			10^{-10}		6	5	8.85	10^{1}	10^{-10}	10^{-6}	OPT.
38.	3	16	10^{-6}		13	8	26.1	10^{1}	10^{-6}	10^{-6}	OPT.
			10^{-10}		13	8	26.1	10^{1}	10^{-6}	10^{-6}	OPT.

Numerical Results for MNA

	n	m	XTOL	max. step	f, J evals.	iters.	$\lVert x^* \rVert_2$	$\lVert f^* \rVert_2$	$\lVert g^* \rVert_2$	est. err.	conv.
39a.	2	3	10^{-6}		4	3	10^{-6}	10^{-1}	10^{-17}	10^{-7}	OPT.
			10^{-10}		4	3	10^{-6}	10^{-1}	10^{-17}	10^{-7}	OPT.
39b.	2	3	10^{-6}		4	3	10^{-7}	10^{-1}	10^{-11}	10^{-7}	OPT
			10^{-10}		4	3	10^{-7}	10^{-1}	10^{-11}	10^{-7}	OPT.
39c.	2	3	10^{-6}		4	3	10^{-7}	10^{-1}	10^{-13}	10^{-7}	OPT.
			10^{-10}		4	3	10^{-7}	10^{-1}	10^{-13}	10^{-7}	OPT.
39d.	2	3	10^{-6}		6	5	10^{-7}	10^{-1}	10^{-17}	10^{-7}	OPT.
			10^{-10}		6	5	10^{-7}	10^{-1}	10^{-17}	10^{-7}	OPT.
39e.	2	3	10^{-6}		8	7	10^{-8}	10^{-1}	10^{-14}	10^{-7}	OPT.
			10^{-10}		8	7	10^{-8}	10^{-1}	10^{-14}	10^{-7}	OPT.
39f.	2	3	10^{-6}		11	10	10^{-9}	10^{-1}	10^{-8}	10^{-7}	*
			10^{-10}		11	10	10^{-9}	10^{-1}	10^{-8}	10^{-7}	*
39g.	2	3	10^{-6}		14	13	10^{-10}	10^{-1}	10^{-9}	10^{-7}	*
			10^{-10}		14	13	10^{-10}	10^{-1}	10^{-9}	10^{-7}	*
40a.	3	4	10^{-6}		4	3	10^{-6}	10^{0}	10^{-15}	10^{-7}	OPT.
			10^{-10}		4	3	10^{-6}	10^{0}	10^{-15}	10^{-7}	OPT.
40b.	3	4	10^{-6}		4	3	10^{-6}	10^{0}	10^{-10}	10^{-7}	OPT.
			10^{-10}		4	3	10^{-6}	10^{0}	10^{-10}	10^{-7}	OPT.
40c.	3	4	10^{-6}		5	4	10^{-7}	10^{0}	10^{-12}	10^{-7}	OPT.
			10^{-10}		5	4	10^{-7}	10^{0}	10^{-12}	10^{-7}	OPT.
40d.	3	4	10^{-6}		6	5	10^{-7}	10^{0}	10^{-16}	10^{-7}	OPT.
			10^{-10}		6	5	10^{-7}	10^{0}	10^{-16}	10^{-7}	OPT.
40e.	3	4	10^{-6}		8	7	10^{-7}	10^{0}	10^{-8}	10^{-7}	*
			10^{-10}		8	7	10^{-7}	10^{0}	10^{-8}	10^{-7}	*
40f.	3	4	10^{-6}		10	9	10^{-8}	10^{0}	10^{-11}	10^{-7}	OPT.
			10^{-10}		10	9	10^{-8}	10^{0}	10^{-11}	10^{-7}	OPT.
40g.	3	4	10^{-6}		13	12	10^{-9}	10^{0}	10^{-6}	10^{-7}	*
			10^{-10}		13	12	10^{-9}	10^{0}	10^{-6}	10^{-7}	*
41a.	5	10	10^{-6}		4	3	10^{-6}	10^{0}	10^{-9}	10^{-7}	*
			10^{-10}		4	3	10^{-6}	10^{0}	10^{-9}	10^{-7}	*
41b.	5	10	10^{-6}		4	3	10^{-6}	10^{0}	10^{-13}	10^{-7}	OPT
			10^{-10}		4	3	10^{-6}	10^{0}	10^{-13}	10^{-7}	OPT.
41c.	5	10	10^{-6}		8	7	10^{-6}	10^{0}	10^{-12}	10^{-7}	OPT
			10^{-10}		8	7	10^{-6}	10^{0}	10^{-12}	10^{-7}	OPT
41d.	5	10	10^{-6}		9	8	10^{-6}	10^{0}	10^{-14}	10^{-7}	OPT.
			10^{-10}		9	8	10^{-6}	10^{0}	10^{-14}	10^{-7}	OPT.
41e.	5	10	10^{-6}		12	11	10^{-7}	10^{0}	10^{-10}	10^{-7}	*
			10^{-10}		12	11	10^{-7}	10^{0}	10^{-10}	10^{-7}	*
41f.	5	10	10^{-6}		14	13	10^{-7}	10^{0}	10^{-11}	10^{-7}	OPT.
			10^{-10}		14	13	10^{-7}	10^{0}	10^{-11}	10^{-7}	OPT.
41g.	5	10	10^{-6}		17	16	10^{-8}	10^{0}	10^{-12}	10^{-7}	OPT.
			10^{-10}		17	16	10^{-8}	10^{0}	10^{-12}	10^{-7}	OPT.

Numerical Results for MNA

	n	m	XTOL	max. step	f, J evals.	iters.	$\|r^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
42a.^{0}	4	24	10^{-6}	2.0	41	35	60.3	10^{-10}	10^{-8}	10^{-20}	*
			10^{-10}	2.0	41	35	60.3	10^{-10}	10^{-8}	10^{-20}	*
42b.^{0}	4	24	10^{-6}		16	8	60.3	10^{-10}	10^{-8}	10^{-21}	*
			10^{-10}		16	8	60.3	10^{-10}	10^{-8}	10^{-21}	*
42c.^{0}	4	24	10^{-6}		6	4	60.3	10^{-11}	10^{-9}	10^{-22}	*
			10^{-10}		6	4	60.3	10^{-11}	10^{-9}	10^{-22}	*
42d.^{0}	4	24	10^{-6}		6	4	60.3	10^{-13}	10^{-11}	10^{-27}	OPT.
			10^{-10}		6	4	60.3	10^{-13}	10^{-11}	10^{-27}	OPT.
43a.^{0}	5	16	10^{-6}	10.0	30	16	54.0	10^{-14}	10^{-11}	10^{-27}	OPT.
			10^{-10}	10.0	30	16	54.0	10^{-14}	10^{-11}	10^{-27}	OPT.
43b.^{0}	5	16	10^{-6}	10.0	17	12	54.0	10^{-13}	10^{-11}	10^{-25}	OPT.
			10^{-10}	10.0	17	12	54.0	10^{-13}	10^{-11}	10^{-25}	OPT.
43c.^{0}	5	16	10^{-6}	10.0	89	45	54.0	10^{-13}	10^{-11}	10^{-25}	OPT.
			10^{-10}	10.0	89	45	54.0	10^{-13}	10^{-11}	10^{-25}	OPT.
43d.^{0}	5	16	10^{-6}	10^{2}	41	21	54.0	10^{-12}	10^{-10}	10^{-27}	OPT.
			10^{-10}	10^{2}	42	22	54.0	10^{-12}	10^{-10}	10^{-27}	*
43e.^{0}	5	16	10^{-6}	10.0	142	73	54.0	10^{-12}	10^{-10}	10^{-23}	OPT.
			10^{-10}	10.0	143	74	54.0	10^{-12}	10^{-10}	10^{-23}	*
43f.^{0}	5	16	10^{-6}	10^{2}	37	19	54.0	10^{-12}	10^{-10}	10^{-25}	OPT.
			10^{-10}	10^{2}	37	19	54.0	10^{-12}	10^{-10}	10^{-25}	OPT.
44a.^{0}	6	6	10^{-6}		143	49	4.03	10^{-12}	10^{-10}	10^{-23}	OPT.
			10^{-10}		144	50	4.03	10^{-12}	10^{-10}	10^{-23}	*
44b.^{0}	6	6	10^{-6}		46	14	3.52	10^{-13}	10^{-11}	10^{-25}	OPT.
			10^{-10}		46	14	3.52	10^{-13}	10^{-11}	10^{-25}	OPT.
44c.^{0}	6	6	10^{-6}		914	364	20.6	10^{-12}	10^{-8}	10^{-23}	OPT.
			10^{-10}		915	365	20.6	10^{-12}	10^{-8}	10^{-23}	*
44d.^{0}	6	6	10^{-6}		915	342	15.3	10^{-12}	10^{-9}	10^{-23}	OPT.
			10^{-10}		916	343	15.3	10^{-12}	10^{-9}	10^{-23}	*
44e.^{0}	6	6	10^{-6}		475	153	9.27	10^{-10}	10^{-7}	10^{-19}	OPT
			10^{-10}		476	154	9.27	10^{-14}	10^{-11}	10^{-28}	OPT.
45a.^{0}	8	8	10^{-6}		186	52	4.06	10^{-16}	10^{-14}	10^{-31}	OPT
			10^{-10}		186	52	4.06	10^{-16}	10^{-14}	10^{-31}	OPT.
45b.^{0}	8	8	10^{-6}		38	15	3.56	10^{-11}	10^{-9}	10^{-21}	*
			10^{-10}		38	15	3.56	10^{-11}	10^{-9}	10^{-21}	*
45c.^{0}	8	8	10^{-6}		1416	578	20.6	10^{-14}	10^{-10}	10^{-28}	OPT
			10^{-10}		1416	578	20.6	10^{-14}	10^{-10}	10^{-28}	OPT.
45d.^{0}	8	8	10^{-6}		1478	586	15.3	10^{-13}	10^{-10}	10^{-25}	OPT.
			10^{-10}		1479	587	15.3	10^{-13}	10^{-10}	10^{-25}	*
45e.^{0}	8	8	10^{-6}		1441	489	9.31	10^{-14}	10^{-11}	10^{-28}	OPT
			10^{-10}		1441	489	9.31	10^{-14}	10^{-11}	10^{-28}	OPT

5.5 Gauss-Newton Methods

5.5.1 Software and Algorithm

The software package LSSOL [Gill et al. (1986a)] is used to solve the linear least-squares subproblem (3.1.3). The linesearch procedure used for the numerical examples in this section, requires both function and gradient information. It is taken from the nonlinear programming code NPSOL [Gill et al. (1979); (1986b)].

5.5.2 Parameters

Parameters in LSSOL were kept at their default values with the following exceptions :

Rank Tolerance	-	varied, see tables
Infinite Bound Size	-	10^{20}

See Gill et al. [1986a] for details concerning the parameters.

In addition, the following parameters are chosen for the linesearch :

η	-	0.5
α_{max}	-	$\min\{(100(1+\|x\|_2)+1)/\|p\|_2, 10^{20}\}$ †

† In some cases the default value α_{max} was too large and overflow occurred during function evaluation in the linesearch. These cases are indicated in the tables by giving the value $\gamma < 100$ such that $\alpha_{max} == \min\{(\gamma(1+\|x\|_2)+1)/\|p\|_2, 10^{20})\}$ that was subsequently used to obtain the results in the column labeled "step fac.".

See, e. g., Gill, Murray, and Wright [1981] for a discussion of the linesearch parameters.

5.5.3 Convergence Criteria

Convergence is judged to have occurred at the kth iterate if either

$$\|f_k\|_2 \le \epsilon_M^{0.9} \tag{5.5.1}$$

or

$$\|g_k\|_2 \le \epsilon_M^{2/3}(1+\|f_k\|_2). \tag{5.5.2}$$

The algorithm is also terminated if there is a negligible change in x,

$$\alpha_k\|p_k\|_2 \le \epsilon_M^{0.9}(1+\|x_k\|_2). \tag{5.5.3}$$

where α_k is the step length determined by the linesearch.

5.5.4 Table Information

Under the label 'conv.', the following notation is used to describe conditions under which the algorithm terminates :

ABS. F	-	(5.5.1)
G	-	(5.5.2)
X	-	(5.5.3)
F LIM.	-	function evaluation limit reached

A superscript ^{0} following a problem number indicates a zero-residual problem.

A superscript ^{L} following a problem number denotes a linear least-squares problem.

Numerical Results for some Gauss-Newton Methods

	n	m	rank tol.	step fac.	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
$1.^0$	2	2	1.49×10^{-8}		31	12	1.41	10^{-17}	10^{-17}	10^{-34}	ABS. F, G
			2.23×10^{-16}		31	12	1.41	10^{-17}	10^{-17}	10^{-34}	ABS. F, G
$2.^0$	2	2	1.49×10^{-8}		180	42	11.4	10^{1}	10^{-7}	10^{1}	X
			2.23×10^{-16}		235	42	11.4	10^{1}	10^{-2}	10^{1}	X
$3.^0$	2	2	1.49×10^{-8}		31	16	7.22	10^{-3}	10^{-7}	10^{-7}	X
			2.23×10^{-16}		42	23	9.11	10^{-16}	10^{-11}	10^{-32}	ABS. F, G
$4.^0$	2	3	1.49×10^{-8}		54	14	10^{6}	10^{-16}	10^{-10}	10^{-32}	ABS. F, X
			2.23×10^{-16}		54	14	10^{6}	10^{-16}	10^{-10}	10^{-32}	ABS. F, X
$5.^0$	2	3	1.49×10^{-8}		8	6	3.04	10^{-14}	10^{-13}	10^{-28}	G
			2.23×10^{-16}		8	6	3.04	10^{-14}	10^{-13}	10^{-28}	G
6.	2	10	1.49×10^{-8}	5.0	(2003)	(294)	.367	10^{1}	10^{2}	10^{-3}	F LIM.
			2.23×10^{-16}	5.0	301	44	.501	10^{2}	10^{4}	10^{1}	X
$7.^0$	3	3	1.49×10^{-8}		13	10	1.00	10^{-24}	10^{-28}	10^{-48}	ABS. F, G
			2.23×10^{-16}		13	10	1.00	10^{-24}	10^{-28}	10^{-48}	ABS. F, G
8.	3	15	1.49×10^{-8}		7	6	2.60	10^{-1}	10^{-11}	10^{-8}	G
			2.23×10^{-16}		7	6	2.60	10^{-1}	10^{-11}	10^{-8}	G
9.	3	15	1.49×10^{-8}		3	2	1.08	10^{-4}	10^{-12}	10^{-14}	G
			2.23×10^{-16}		3	2	1.08	10^{-4}	10^{-12}	10^{-14}	G
10.	3	16	1.49×10^{-8}		112	29	10^{4}	10^{1}	10^{0}	10^{1}	X
			2.23×10^{-16}		30	10	10^{4}	10^{1}	10^{-3}	10^{-6}	X
$11.^0$	3	10	1.49×10^{-8}		(3000)	(1300)	252.	10^{-1}	10^{3}	10^{-2}	F LIM.
			2.23×10^{-16}		(3000)	(999)	308.	10^{-1}	10^{-1}	10^{-2}	F LIM.
$12.^0$	3	10	1.49×10^{-8}		7	6	10.1	10^{-17}	10^{-17}	10^{-33}	ABS. F, G
			2.23×10^{-16}		7	6	10.1	10^{-17}	10^{-17}	10^{-33}	ABS. F, G
$13.^0$	4	4	1.49×10^{-8}		16	15	10^{-4}	10^{-8}	10^{-11}	10^{-16}	G
			2.23×10^{-16}		16	15	10^{-4}	10^{-8}	10^{-11}	10^{-16}	G
$14.^0$	4	6	1.49×10^{-8}		96	42	2.00	0.00	0.00	0.00	ABS. F, G
			2.23×10^{-16}		96	42	2.00	0.00	0.00	0.00	ABS. F, G
15.	4	11	1.49×10^{-8}		43	35	.328	10^{-2}	10^{-11}	10^{-9}	G
			2.23×10^{-16}		43	35	.328	10^{-2}	10^{-11}	10^{-9}	G
16.	4	20	1.49×10^{-8}		3651	1765	17.6	10^{2}	10^{-8}	10^{-8}	X
			2.23×10^{-16}		3651	1765	17.6	10^{2}	10^{-8}	10^{-8}	X
17.	5	33	1.49×10^{-8}		13	9	2.46	10^{-2}	10^{-11}	10^{-11}	G
			2.23×10^{-16}		13	9	2.46	10^{-2}	10^{-11}	10^{-11}	G
$18.^0$	6	13	1.49×10^{-8}	10.0	(6001)	(770)	2.99	10^{0}	10^{0}	10^{0}	F LIM.
			2.23×10^{-16}		(6001)	(856)	52.5	10^{0}	10^{0}	10^{0}	F LIM.
19.	11	65	1.49×10^{-8}		24	16	9.38	10^{-1}	10^{-11}	10^{-8}	G
			2.23×10^{-16}		24	16	9.38	10^{-1}	10^{-11}	10^{-8}	G

Numerical Results for some Gauss-Newton Methods

	n	m	rank tol.	step fac.	f, J evals.	iters.	$\lVert x^* \rVert_2$	$\lVert f^* \rVert_2$	$\lVert g^* \rVert_2$	est. err.	conv.
20a.	6	31	1.49×10^{-8}		12	11	2.44	10^{-2}	10^{-11}	10^{-10}	G
			2.23×10^{-16}		12	11	2.44	10^{-2}	10^{-11}	10^{-10}	G
20b.	9	31	1.49×10^{-8}		6	5	6.06	10^{-3}	10^{-11}	10^{-13}	G
			2.23×10^{-16}		6	5	6.06	10^{-3}	10^{-11}	10^{-13}	G
20c.	12	31	1.49×10^{-8}		6	5	16.6	10^{-5}	10^{-14}	10^{-16}	G
			2.23×10^{-16}		6	5	16.6	10^{-5}	10^{-14}	10^{-16}	G
20d.	20	31	1.49×10^{-8}		6	5	1.07	10^{-8}	10^{-13}	10^{-15}	G
			2.23×10^{-16}		6	5	247.	10^{-10}	10^{-12}	10^{-25}	G
21a.^{0}	10	10	1.49×10^{-8}		31	12	3.16	10^{-16}	10^{-14}	10^{-31}	ABS. F, G
			2.23×10^{-16}		31	12	3.16	10^{-16}	10^{-14}	10^{-31}	ABS. F, G
21b.^{0}	20	20	1.49×10^{-8}		31	12	4.47	10^{-15}	10^{-14}	10^{-30}	ABS. F, G
			2.23×10^{-16}		31	12	4.47	10^{-15}	10^{-14}	10^{-30}	ABS. F, G
22a.^{0}	12	12	1.49×10^{-8}		16	15	10^{-4}	10^{-8}	10^{-11}	10^{-16}	G
			2.23×10^{-16}		16	15	10^{-4}	10^{-8}	10^{-11}	10^{-16}	G
22b.^{0}	20	20	1.49×10^{-8}		16	15	10^{-4}	10^{-8}	10^{-11}	10^{-15}	G
			2.23×10^{-16}		16	15	10^{-4}	10^{-8}	10^{-11}	10^{-15}	G
23a.	4	5	1.49×10^{-8}		90	45	.500	10^{-3}	10^{-13}	10^{-10}	G
			2.23×10^{-16}		90	45	.500	10^{-3}	10^{-13}	10^{-10}	G
23b.	10	11	1.49×10^{-8}		274	122	.500	10^{-2}	10^{-12}	10^{-11}	G
			2.23×10^{-16}		274	122	.500	10^{-2}	10^{-12}	10^{-11}	G
24a.	4	8	1.49×10^{-8}		1043	302	.759	10^{-3}	10^{-12}	10^{-11}	G
			2.23×10^{-16}		1043	302	.759	10^{-3}	10^{-12}	10^{-11}	G
24b.	10	20	1.49×10^{-8}		(10003)	(2556)	.598	10^{-2}	10^{-3}	10^{-7}	F LIM.
			2.23×10^{-16}		(10003)	(2556)	.598	10^{-2}	10^{-3}	10^{-7}	F LIM.
25a.^{0}	10	12	1.49×10^{-8}		11	10	3.16	10^{-15}	10^{-14}	10^{-30}	ABS. F, G
			2.23×10^{-16}		11	10	3.16	10^{-15}	10^{-14}	10^{-30}	ABS. F, G
25b.^{0}	20	22	1.49×10^{-8}		13	12	4.47	10^{-15}	10^{-13}	10^{-30}	ABS. F, G
			2.23×10^{-16}		13	12	4.47	10^{-15}	10^{-13}	10^{-30}	ABS. F, G
26a.^{0}	10	10	1.49×10^{-8}		16	8	.306	10^{-11}	10^{-11}	10^{-22}	G
			2.23×10^{-16}		16	8	.306	10^{-11}	10^{-11}	10^{-22}	G
26b.^{0}	20	20	1.49×10^{-8}		20	9	.208	10^{-14}	10^{-14}	10^{-28}	G
			2.23×10^{-16}		20	9	.208	10^{-14}	10^{-14}	10^{-28}	G
27a.^{0}	10	10	1.49×10^{-8}		21	7	3.18	10^{-15}	10^{-14}	10^{-29}	ABS. F, G
			2.23×10^{-16}		21	7	3.18	10^{-15}	10^{-14}	10^{-29}	ABS. F, G
27b.^{0}	20	20	1.49×10^{-8}	10.0	31	7	4.47	10^{-14}	10^{-13}	10^{-27}	G
			2.23×10^{-16}	10.0	31	7	4.47	10^{-14}	10^{-13}	10^{-27}	G
28a.^{0}	10	10	1.49×10^{-8}		4	3	.412	10^{-15}	10^{-16}	10^{-31}	ABS. F, G
			2.23×10^{-16}		4	3	.412	10^{-15}	10^{-16}	10^{-31}	ABS. F, G
28b.^{0}	20	20	1.49×10^{-8}		4	3	.571	10^{-16}	10^{-16}	10^{-32}	ABS. F, G
			2.23×10^{-16}		4	3	.571	10^{-16}	10^{-16}	10^{-32}	ABS. F, G

Numerical Results for some Gauss-Newton Methods

	n	m	rank tol.	step fac.	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
29a.[0]	10	10	1.49×10^{-8}		4	3	.412	10^{-14}	10^{-14}	10^{-29}	ABS. F, G
			2.23×10^{-16}		4	3	.412	10^{-14}	10^{-14}	10^{-29}	ABS. F, G
29b.[0]	20	20	1.49×10^{-8}		4	3	.571	10^{-14}	10^{-14}	10^{-28}	G
			2.23×10^{-16}		4	3	.571	10^{-14}	10^{-14}	10^{-28}	G
30a.[0]	10	10	1.49×10^{-8}		6	5	2.05	10^{-16}	10^{-15}	10^{-31}	ABS. F, G
			2.23×10^{-16}		6	5	2.05	10^{-16}	10^{-15}	10^{-31}	ABS. F, G
30b.[0]	20	20	1.49×10^{-8}		6	5	3.04	10^{-15}	10^{-15}	10^{-31}	ABS. F, G
			2.23×10^{-16}		6	5	3.04	10^{-15}	10^{-15}	10^{-31}	ABS. F, G
31a.[0]	10	10	1.49×10^{-8}		7	6	1.80	10^{-15}	10^{-15}	10^{-31}	ABS. F, G
			2.23×10^{-16}		7	6	1.80	10^{-15}	10^{-15}	10^{-31}	ABS. F, G
31b.[0]	20	20	1.49×10^{-8}		7	6	2.66	10^{-15}	10^{-15}	10^{-31}	ABS. F, G
			2.23×10^{-16}		7	6	2.66	10^{-15}	10^{-15}	10^{-31}	ABS. F, G
32.[L]	10	20	1.49×10^{-8}		2	1	3.16	10^{0}	10^{-14}	0.00	G
			2.23×10^{-16}		2	1	3.16	10^{0}	10^{-14}	0.00	G
33.[L]	10	20	1.49×10^{-8}		3	2	5.40	10^{0}	10^{-10}	10^{-6}	G, X
			2.23×10^{-16}		8	8	10^{13}	10^{0}	10^{2}	10^{-3}	$g^Tp \geq 0$
34.[L]	10	20	1.49×10^{-8}		3	2	4.90	10^{0}	10^{-11}	10^{-6}	G, X
			2.23×10^{-16}		3	2	4.90	10^{0}	10^{-11}	10^{-6}	G, X
35a.	8	8	1.49×10^{-8}		3053	386	1.61	10^{-1}	10^{-1}	10^{-3}	X
			2.23×10^{-16}		(8003)	(1012)	1.60	10^{-1}	10^{0}	10^{-2}	F LIM.
35b.[0]	9	9	1.49×10^{-8}		148	29	1.73	10^{-14}	10^{-14}	10^{-29}	ABS. F, G
			2.23×10^{-16}		249	38	1.70	10^{-1}	10^{0}	10^{-2}	X
35c.	10	10	1.49×10^{-8}		(10006)	(1353)	1.79	10^{-1}	10^{0}	10^{-2}	F LIM.
			2.23×10^{-16}		232	31	1.79	10^{-1}	10^{0}	10^{-2}	X
36a.[0]	4	4	1.49×10^{-8}		2885	490	18.8	10^{-6}	10^{-12}	10^{-12}	G
			2.23×10^{-16}		36	21	50.0	10^{-15}	10^{-14}	10^{-31}	ABS. F, G
36b.[0]	9	9	1.49×10^{-8}		683	128	18.8	10^{-6}	10^{-12}	10^{-12}	G
			2.23×10^{-16}		36	21	50.0	10^{-15}	10^{-14}	10^{-31}	ABS. F, G
36c.[0]	9	9	1.49×10^{-8}		20	19	1.73	10^{-11}	10^{-11}	10^{-22}	G
			2.23×10^{-16}		20	19	1.73	10^{-11}	10^{-11}	10^{-22}	G
36d.[0]	9	9	1.49×10^{-8}		74	29	19.1	10^{-6}	10^{-11}	10^{-12}	G
			2.23×10^{-16}		(9001)	(1190)	345.	10^{-6}	10^{-5}	10^{-13}	F LIM.
37.	2	16	1.49×10^{-8}		39	38	8.85	10^{1}	10^{-8}	10^{-6}	X
			2.23×10^{-16}		39	38	8.85	10^{1}	10^{-8}	10^{-6}	X
38.	3	16	1.49×10^{-8}		58	56	26.1	10^{1}	10^{-10}	10^{-6}	G
			2.23×10^{-16}		58	56	26.1	10^{1}	10^{-10}	10^{-6}	G

Numerical Results for some Gauss-Newton Methods

	n	m	rank tol.	step fac.	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
39a.	2	3	1.49×10^{-8}		8	7	10^{-6}	10^{-1}	10^{-11}	10^{-7}	G
			2.23×10^{-16}		8	7	10^{-6}	10^{-1}	10^{-11}	10^{-7}	G
39b.	2	3	1.49×10^{-8}		32	31	10^{-7}	10^{-1}	10^{-11}	10^{-7}	G
			2.23×10^{-16}		32	31	10^{-7}	10^{-1}	10^{-11}	10^{-7}	G
39c.	2	3	1.49×10^{-8}		23	14	10^{-7}	10^{-1}	10^{-12}	10^{-7}	G
			2.23×10^{-16}		23	14	10^{-7}	10^{-1}	10^{-12}	10^{-7}	G
39d.	2	3	1.49×10^{-8}		681	333	10^{-7}	10^{-1}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		681	333	10^{-7}	10^{-1}	10^{-10}	10^{-7}	G
39e.	2	3	1.49×10^{-8}		(2001)	(925)	10^{-7}	10^{-1}	10^{-9}	10^{-7}	F LIM
			2.23×10^{-16}		(2001)	(925)	10^{-7}	10^{-1}	10^{-9}	10^{-7}	F LIM.
39f.	2	3	1.49×10^{-8}		(2000)	(873)	10^{-9}	10^{-1}	10^{-8}	10^{-7}	F LIM.
			2.23×10^{-16}		(2000)	(873)	10^{-9}	10^{-1}	10^{-8}	10^{-7}	F LIM.
39g.	2	3	1.49×10^{-8}		(2001)	(607)	10^{-8}	10^{-1}	10^{-5}	10^{-7}	F LIM.
			2.23×10^{-16}		(2001)	(607)	10^{-8}	10^{-1}	10^{-5}	10^{-7}	F LIM.
40a.	3	4	1.49×10^{-8}		13	12	10^{-6}	10^{0}	10^{-11}	10^{-7}	G
			2.23×10^{-16}		13	12	10^{-6}	10^{0}	10^{-11}	10^{-7}	G
40b.	3	4	1.49×10^{-8}		16	10	10^{-6}	10^{0}	10^{-12}	10^{-7}	G
			2.23×10^{-16}		16	10	10^{-6}	10^{0}	10^{-12}	10^{-7}	G
40c.	3	4	1.49×10^{-8}		380	188	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		380	188	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
40d.	3	4	1.49×10^{-8}		781	345	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		781	345	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
40e.	3	4	1.49×10^{-8}		(3000)	(690)	10^{-2}	10^{0}	10^{-1}	10^{-3}	F LIM.
			2.23×10^{-16}		(3000)	(690)	10^{-2}	10^{0}	10^{-1}	10^{-3}	F LIM.
40f.	3	4	1.49×10^{-8}		(3002)	(630)	.101	10^{0}	10^{1}	10^{-1}	F LIM.
			2.23×10^{-16}		(3002)	(630)	.101	10^{0}	10^{1}	10^{-1}	F LIM.
40g.	3	4	1.49×10^{-8}		(3001)	(607)	.110	10^{1}	10^{3}	10^{1}	F LIM.
			2.23×10^{-16}		(3001)	(607)	.110	10^{1}	10^{3}	10^{1}	F LIM.
41a.	5	10	1.49×10^{-8}		5	4	10^{-6}	10^{0}	10^{-13}	10^{-7}	G
			2.23×10^{-16}		5	4	10^{-6}	10^{0}	10^{-13}	10^{-7}	G
41b.	5	10	1.49×10^{-8}		6	5	10^{-6}	10^{0}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		6	5	10^{-6}	10^{0}	10^{-10}	10^{-7}	G
41c.	5	10	1.49×10^{-8}		12	11	10^{-6}	10^{0}	10^{-11}	10^{-7}	G
			2.23×10^{-16}		12	11	10^{-6}	10^{0}	10^{-11}	10^{-7}	G
41d.	5	10	1.49×10^{-8}		30	20	10^{-6}	10^{0}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		30	20	10^{-6}	10^{0}	10^{-10}	10^{-7}	G
41e.	5	10	1.49×10^{-8}		222	110	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		222	110	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
41f.	5	10	1.49×10^{-8}		933	414	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		933	414	10^{-7}	10^{0}	10^{-10}	10^{-7}	G
41g.	5	10	1.49×10^{-8}		3285	1065	10^{-8}	10^{0}	10^{-10}	10^{-7}	G
			2.23×10^{-16}		3285	1065	10^{-8}	10^{0}	10^{-10}	10^{-7}	G

Numerical Results for some Gauss-Newton Methods

	n	m	rank tol.	step fac.	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
42a.[0]	4	24	1.49×10^{-8}	0.1	51	37	60.8	10^{-13}	10^{-11}	10^{-26}	G, X
			2.23×10^{-16}	0.1	51	37	60.8	10^{-13}	10^{-11}	10^{-26}	G, X
42b.[0]	4	24	1.49×10^{-8}	0.1	611	316	61.9	10^{-13}	10^{-10}	10^{-25}	X
			2.23×10^{-16}	0.1	(4002)	(1577)	10^{8}	10^{2}	10^{11}	10^{5}	F LIM.
42c.[0]	4	24	1.49×10^{-8}	0.1	27	24	60.3	10^{-13}	10^{-11}	10^{-26}	G
			2.23×10^{-16}	0.1	27	24	60.3	10^{-13}	10^{-11}	10^{-26}	G
42d.[0]	4	24	1.49×10^{-8}	0.1	24	23	60.3	10^{-13}	10^{-11}	10^{-26}	G, X
			2.23×10^{-16}	0.1	24	23	60.3	10^{-13}	10^{-11}	10^{-26}	G, X
43a.[0]	5	16	1.49×10^{-8}	1.0	23	14	54.0	10^{-14}	10^{-12}	10^{-28}	G
			2.23×10^{-16}	1.0	23	14	54.0	10^{-14}	10^{-12}	10^{-28}	G
43b.[0]	5	16	1.49×10^{-8}	0.1	15	13	53.6	10^{-1}	10^{-7}	10^{-2}	X
			2.23×10^{-16}	0.1	20	14	54.0	10^{-14}	10^{-12}	10^{-28}	G
43c.[0]	5	16	1.49×10^{-8}	1.0	24	14	54.0	10^{-14}	10^{-11}	10^{-27}	G
			2.23×10^{-16}	1.0	24	14	54.0	10^{-14}	10^{-11}	10^{-27}	G
43d.[0]	5	16	1.49×10^{-8}	1.0	18	10	54.0	10^{-14}	10^{-11}	10^{-27}	G
			2.23×10^{-16}	1.0	18	10	54.0	10^{-14}	10^{-11}	10^{-27}	G
43e.[0]	5	16	1.49×10^{-8}	1.0	31	17	54.0	10^{-14}	10^{-12}	10^{-28}	G
			2.23×10^{-16}	1.0	31	17	54.0	10^{-14}	10^{-12}	10^{-28}	G
43f.[0]	5	16	1.49×10^{-8}		22	15	54.0	10^{-14}	10^{-12}	10^{-28}	G
			2.23×10^{-16}		22	15	54.0	10^{-14}	10^{-12}	10^{-28}	G
44a.[0]	6	6	1.49×10^{-8}		171	45	4.06	10^{-13}	10^{-11}	10^{-25}	G
			2.23×10^{-16}		171	45	4.06	10^{-13}	10^{-11}	10^{-25}	G
44b.[0]	6	6	1.49×10^{-8}		5	4	3.52	10^{-15}	10^{-13}	10^{-29}	ABS. F, G
			2.23×10^{-16}		5	4	3.52	10^{-15}	10^{-13}	10^{-29}	ABS. F, G
44c.[0]	6	6	1.49×10^{-8}		55	17	20.6	10^{-14}	10^{-10}	10^{-28}	X
			2.23×10^{-16}		55	17	20.6	10^{-14}	10^{-10}	10^{-28}	X
44d.[0]	6	6	1.49×10^{-8}		35	15	15.3	10^{-14}	10^{-11}	10^{-29}	ABS. F, G
			2.23×10^{-16}		35	15	15.3	10^{-14}	10^{-11}	10^{-29}	ABS. F, G
44e.[0]	6	6	1.49×10^{-8}		42	18	9.27	10^{-14}	10^{-12}	10^{-28}	ABS. F, G, X
			2.23×10^{-16}		42	18	9.27	10^{-14}	10^{-12}	10^{-28}	ABS. F, G, X
45a.[0]	8	8	1.49×10^{-8}		171	45	4.06	10^{-13}	10^{-11}	10^{-25}	G
			2.23×10^{-16}		171	45	4.06	10^{-13}	10^{-11}	10^{-25}	G
45b.[0]	8	8	1.49×10^{-8}		5	4	3.56	10^{-15}	10^{-13}	10^{-29}	ABS. F, G
			2.23×10^{-16}		5	4	3.56	10^{-15}	10^{-13}	10^{-29}	ABS. F, G
45c.[0]	8	8	1.49×10^{-8}		41	17	20.6	10^{-11}	10^{-11}	10^{-29}	ABS. F, X
			2.23×10^{-16}		41	17	20.6	10^{-14}	10^{-11}	10^{-29}	ABS. F, X
45d.[0]	8	8	1.49×10^{-8}		35	15	15.3	10^{-16}	10^{-13}	10^{-31}	ABS. F, G
			2.23×10^{-16}		35	15	15.3	10^{-16}	10^{-13}	10^{-31}	ABS. F, G
45e.[0]	8	8	1.49×10^{-8}		42	18	9.31	10^{-15}	10^{-12}	10^{-29}	ABS. F, G, X
			2.23×10^{-16}		42	18	9.31	10^{-15}	10^{-12}	10^{-29}	ABS. F, G, X

5.6 Levenberg-Marquardt Method

(MINPACK LMDER)

5.6.1 Software and Algorithm

The results were obtained using the MINPACK subroutine **LMDER**, which implements a Levenberg-Marquardt method using exact derivative information. A subproblem of the form

$$\min_{p \in \Re^n} Q_k(p) \equiv 2g_k^T p + p^T J_k^T J_k p$$

$$\text{subject to} \quad \|D_k p\|_2 \leq \delta_k$$

is solved at each iteration for the step p_k to the next iterate, where D_k is a diagonal scaling matrix.

5.6.2 Parameters

The results were obtained using the MINPACK subroutine **LMDER**, with the following input parameters :

XTOL	-	varied, see tables	accuracy in x
FTOL	-	varied, see tables	accuracy in sum of squares
GTOL	-	0.00	gradient norm tolerance
MAXFEV	-	min $\{9999, 1000 * n\}$	function evaluation limit
MODE	-	1	specifies internal scaling
FACTOR	-	100. (default)	initial step magnification

† In some cases the default **FACTOR** = 100.0 was too large and overflow occurred during function evaluation. These cases are indicated in the table by giving the lower value of **FACTOR** that was subsequently used to obtain the results.

For details about these parameters, Moré, Garbow, and Hillstrom [1980].

5.6.3 Convergence Criteria

The following quantities will be used in describing the convergence criteria :

residual vector	:	$f(x_k)$
ith residual gradient	:	$\nabla\phi_i(x_k)$
Jacobian matrix	:	$J(x_k)$
objective function	:	$\mathcal{F}(x_k) \equiv f(x_k)^T f(x_k)$
objective gradient	:	$g_k = \nabla\mathcal{F}(x_k) \equiv 2J(x_k)^T f(x_k)$
current step	:	p_k, the minimizer of the subproblem
predicted reduction	:	$\rho_r = \dfrac{\Vert f_k\Vert_2 - \Vert f_k + J_k p_k\Vert_2}{\Vert f_k\Vert_2} = \dfrac{-Q_k(p_k)}{\Vert f_k\Vert_2}$
actual reduction	:	$\rho_A = \dfrac{\Vert f(x_k)\Vert_2 - \Vert f(x_k + p_k)\Vert_2}{\Vert f_k\Vert_2} = \dfrac{\mathcal{F}(x_k) - \mathcal{F}(x_k + p_k)}{\Vert f_k\Vert_2}$

Criteria for termination of **LMDER** at x_k are as follows:

• *$\mathcal{F}$ convergence.* Both actual and predicted reductions in the sum of squares are at most **FTOL**.

$$|\rho_A| \leq \texttt{FTOL} \quad \text{and} \quad \rho_r \leq \texttt{FTOL} \quad \text{and} \quad \rho_A \leq 2\rho_r \tag{5.6.1}$$

This attempts to guarantee that

$$\Vert f_k\Vert_2 \leq (1 + \texttt{FTOL})\Vert f(x^*)\Vert_2.$$

• *x convergence.* Relative error between two consecutive iterates is at most **XTOL**.

$$\delta_{k+1} \leq \texttt{XTOL}\Vert x_k + p_k\Vert_2 \tag{5.6.2}$$

This attempts to guarantee that

$$\Vert D_k(x_k - x^*)\Vert_2 \leq \texttt{XTOL}\Vert D_k(x^*)\Vert_2.$$

• The cosine of the angle between f_k and any column of J_k is at most GTOL in absolute value.

$$\max_{1\leq i\leq m} \frac{|\nabla\phi_i(x_k)^T f_k|}{\Vert\nabla\phi_i(x_k)\Vert_2\Vert f_k\Vert_2} \leq \texttt{GTOL} \tag{5.6.3}$$

This approximates the necessary condition $g(x_k) = 0$.

• FTOL is too small. No further reduction in the sum of squares is possible.

$$|\rho_A| \leq \epsilon_M \quad \text{and} \quad \rho_r \leq \epsilon_M \quad \text{and} \quad \rho_A \leq 2\rho_r \tag{5.6.4}$$

• XTOL is too small. No further improvement in the approximate solution x_k is possible.

$$\delta_{k+1} \leq \epsilon_M \|x_k + p_k\|_2 \tag{5.6.5}$$

• GTOL is too small. f_k is orthogonal to the columns of J_k to machine precision.

$$\max_{1 \leq i \leq n} \frac{\left| \nabla \phi_i(x_k)^T f_k \right|}{\|\nabla \phi_i(x_k)\|_2 \|f_k\|_2} \leq \epsilon_M \tag{5.6.6}$$

Except for test (5.6.3), tests for convergence are performed only when

$$\rho_A < 0.0001 \rho_P. \tag{5.6.7}$$

The convergence criteria are described in more detail in Moré, Garbow, and Hillstrom [1980]. The following abbreviations are used in the tables to describe the conditions under which the algorithm terminates:

F	-	(5.6.1) and (5.6.7)
X	-	(5.6.2) and (5.6.7)
X, F	-	(5.6.1) and (5.6.2) and (5.6.7)
G	-	(5.6.6) and (5.6.7)
F LIM.	-	function evaluation limit reached

Numerical Results for LMDER

	n	m	TOL	max. step	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
1.^{0}	2	2	10^{-8}		22	16	1.41	10^{-16}	10^{-15}	10^{-32}	X
			10^{-12}		22	16	1.41	10^{-16}	10^{-15}	10^{-32}	X
2.^{0}	2	2	10^{-8}		14	8	11.4	10^{1}	10^{-3}	10^{1}	REL. F
			10^{-12}		21	13	11.4	10^{1}	10^{-5}	10^{1}	REL. F
3.^{0}	2	2	10^{-8}		19	17	9.11	10^{-16}	10^{-11}	10^{-32}	X
			10^{-12}		19	17	9.11	10^{-16}	10^{-11}	10^{-32}	X
4.^{0}	2	3	10^{-8}		40	39	10^{6}	10^{-3}	10^{-2}	10^{-5}	X
			10^{-12}		54	53	10^{6}	10^{-7}	10^{-7}	10^{-14}	X
5.^{0}	2	3	10^{-8}		9	7	3.04	10^{-16}	10^{-15}	10^{-32}	X
			10^{-12}		10	8	3.04	10^{-16}	10^{-15}	10^{-32}	X
6.	2	10	10^{-8}		21	12	.365	10^{1}	10^{-2}	10^{-6}	REL. F
			10^{-12}		28	16	.365	10^{1}	10^{-4}	10^{-6}	REL. F
7.^{0}	3	3	10^{-8}		11	8	1.00	10^{-16}	10^{-15}	10^{-32}	X
			10^{-12}		12	9	1.00	10^{-32}	10^{-31}	10^{-64}	X
8.	3	15	10^{-8}		6	5	2.60	10^{-1}	10^{-9}	10^{-8}	REL. F
			10^{-12}		7	6	2.60	10^{-1}	10^{-11}	10^{-8}	REL. F
9.	3	15	10^{-8}		4	3	1.08	10^{-4}	10^{-16}	10^{-14}	X
			10^{-12}		5	4	1.08	10^{-4}	10^{-16}	10^{-14}	X, REL. F
10.	3	16	10^{-8}		126	116	10^{4}	10^{1}	10^{0}	10^{-6}	X, REL. F
			10^{-12}		126	116	10^{4}	10^{1}	10^{0}	10^{-6}	X
11.^{0}	3	10	10^{-8}		(3000)	(2956)	239.	10^{-2}	10^{18}	10^{-5}	F LIM.
			10^{-12}		(3000)	(2956)	239.	10^{-2}	10^{18}	10^{-5}	F LIM.
12.^{0}	3	10	10^{-8}		7	6	10.1	10^{-16}	10^{-16}	10^{-32}	X
			10^{-12}		8	7	10.1	10^{-16}	10^{-16}	10^{-32}	X
13.^{0}	4	4	10^{-8}		65	60	10^{-17}	10^{-34}	10^{-50}	10^{-67}	σ
			10^{-12}		65	60	10^{-17}	10^{-34}	10^{-50}	10^{-67}	σ
14.^{0}	4	6	10^{-8}		70	64	2.00	0.00	0.00	0.00	X
			10^{-12}		70	64	2.00	0.00	0.00	0.00	X
15.	4	11	10^{-8}		18	16	.328	10^{-2}	10^{-7}	10^{-9}	REL. F
			10^{-12}		28	26	.328	10^{-2}	10^{-9}	10^{-9}	X
16.	4	20	10^{-8}		264	245	17.6	10^{2}	10^{0}	10^{-7}	REL. F
			10^{-12}		356	329	17.6	10^{2}	10^{-2}	10^{-8}	REL. F
17.	5	33	10^{-8}		18	15	2.46	10^{-2}	10^{-8}	10^{-11}	REL. F
			10^{-12}		19	16	2.46	10^{-2}	10^{-9}	10^{-11}	REL. F
18.^{0}	6	13	10^{-8}		46	32	12.3	10^{-16}	10^{-15}	10^{-31}	X
			10^{-12}		46	32	12.3	10^{-16}	10^{-15}	10^{-31}	X
19.	11	65	10^{-8}		17	13	9.38	10^{-1}	10^{-7}	10^{-8}	REL. F
			10^{-12}		19	15	9.38	10^{-1}	10^{-9}	10^{-8}	REL. F

Numerical Results for LMDER

	n	m	TOL	max. step	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
20a.	6	31	10^{-8}		8	7	2.44	10^{-2}	10^{-7}	10^{-10}	REL. F
			10^{-12}		10	9	2.44	10^{-2}	10^{-9}	10^{-10}	X, REL. F
20b.	9	31	10^{-8}		9	8	6.06	10^{-3}	10^{-13}	10^{-13}	X
			10^{-12}		10	9	6.06	10^{-3}	10^{-13}	10^{-13}	X, REL. F
20c.	12	31	10^{-8}		10	9	16.6	10^{-5}	10^{-13}	10^{-16}	X
			10^{-12}		12	10	16.6	10^{-5}	10^{-13}	10^{-16}	X
20d.	20	31	10^{-8}		18	14	247.	10^{-10}	10^{-12}	10^{-24}	X
			10^{-12}		23	15	247.	10^{-10}	10^{-12}	10^{-25}	X
21a.[0]	10	10	10^{-8}		22	16	3.16	10^{-16}	10^{-14}	10^{-31}	X
			10^{-12}		22	16	3.16	10^{-16}	10^{-14}	10^{-31}	X
21b.[0]	20	20	10^{-8}		22	16	4.47	10^{-16}	10^{-14}	10^{-31}	X
			10^{-12}		22	16	4.47	10^{-16}	10^{-14}	10^{-31}	X
22a.[0]	12	12	10^{-8}		72	63	10^{-17}	10^{-34}	10^{-51}	10^{-68}	G
			10^{-12}		72	63	10^{-17}	10^{-34}	10^{-51}	10^{-68}	G
22b.[0]	20	20	10^{-8}		69	60	10^{-17}	10^{-33}	10^{-49}	10^{-66}	X
			10^{-12}		69	60	10^{-17}	10^{-33}	10^{-49}	10^{-66}	G
23a.	4	5	10^{-8}		34	23	.500	10^{-3}	10^{-9}	10^{-10}	REL. F
			10^{-12}		44	28	.500	10^{-3}	10^{-11}	10^{-10}	REL. F
23b.	10	11	10^{-8}		84	67	.500	10^{-2}	10^{-8}	10^{-11}	REL. F
			10^{-12}		104	82	.500	10^{-2}	10^{-10}	10^{-11}	REL. F
24a.	4	8	10^{-8}		151	113	.759	10^{-3}	10^{-8}	10^{-11}	REL. F
			10^{-12}		156	116	.759	10^{-3}	10^{-11}	10^{-11}	REL. F
24b.	10	20	10^{-8}		80	62	.598	10^{-2}	10^{-7}	10^{-9}	REL. F
			10^{-12}		88	67	.598	10^{-2}	10^{-10}	10^{-9}	REL. F
25a.[0]	10	12	10^{-8}		11	10	3.16	10^{-15}	10^{-14}	10^{-30}	X
			10^{-12}		12	11	3.16	10^{-16}	10^{-14}	10^{-31}	X
25b.[0]	20	22	10^{-8}		13	12	4.47	10^{-15}	10^{-13}	10^{-30}	X
			10^{-12}		14	13	4.47	10^{-15}	10^{-14}	10^{-31}	X
26a.[0]	10	10	10^{-8}		28	16	.328	10^{-2}	10^{-7}	10^{-5}	REL. F
			10^{-12}		37	21	.328	10^{-2}	10^{-9}	10^{-5}	REL. F
26b.[0]	20	20	10^{-8}		57	40	.228	10^{-3}	10^{-8}	10^{-6}	REL. F
			10^{-12}		71	45	.228	10^{-3}	10^{-10}	10^{-6}	REL. F
27a.[0]	10	10	10^{-8}		15	13	3.18	10^{-15}	10^{-14}	10^{-31}	X
			10^{-12}		15	13	3.18	10^{-15}	10^{-14}	10^{-31}	X
27b.[0]	20	20	10^{-8}		5	2	19.7	10^{0}	10^{-12}	10^{0}	REL. F
			10^{-12}		18	15	4.47	10^{-14}	10^{-13}	10^{-28}	X
28a.[0]	10	10	10^{-8}		5	4	.412	10^{-17}	10^{-16}	10^{-33}	X
			10^{-12}		5	4	.412	10^{-17}	10^{-16}	10^{-33}	X
28b.[0]	20	20	10^{-8}		5	4	.571	10^{-17}	10^{-16}	10^{-33}	X
			10^{-12}		5	4	.571	10^{-17}	10^{-16}	10^{-32}	X

Numerical Results for LMDER

	n	m	TOL	max. step	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
29a.[0]	10	10	10^{-8}		5	4	.412	10^{-17}	10^{-16}	10^{-33}	X
			10^{-12}		5	4	.412	10^{-17}	10^{-16}	10^{-33}	X
29b.[0]	20	20	10^{-8}		5	4	.571	10^{-16}	10^{-16}	10^{-32}	X
			10^{-12}		5	4	.571	10^{-16}	10^{-16}	10^{-32}	X
30a.[0]	10	10	10^{-8}		6	5	2.05	10^{-16}	10^{-15}	10^{-31}	X
			10^{-12}		7	6	2.05	10^{-16}	10^{-15}	10^{-31}	X
30b.[0]	20	20	10^{-8}		6	5	3.04	10^{-15}	10^{-14}	10^{-30}	X
			10^{-12}		7	6	3.04	10^{-15}	10^{-14}	10^{-30}	X
31a.[0]	10	10	10^{-8}		7	6	1.80	10^{-16}	10^{-15}	10^{-31}	X
			10^{-12}		8	7	1.80	10^{-16}	10^{-15}	10^{-31}	X
31b.[0]	20	20	10^{-8}		7	6	2.66	10^{-15}	10^{-14}	10^{-30}	X
			10^{-12}		8	7	2.66	10^{-15}	10^{-14}	10^{-30}	X
32.[L]	10	20	10^{-8}		3	2	3.16	10^{0}	10^{-14}	10^{-17}	X, REL. F
			10^{-12}		3	2	3.16	10^{0}	10^{-14}	10^{-17}	X, REL. F
33.[L]	10	20	10^{-8}		3	2	470.	10^{0}	10^{-7}	10^{-6}	REL. F
			10^{-12}		8	2	470.	10^{0}	10^{-7}	10^{-6}	REL. F
34.[L]	10	20	10^{-8}		3	2	381.	10^{0}	10^{-8}	10^{-6}	REL. F
			10^{-12}		7	3	428.	10^{0}	10^{-9}	10^{-6}	REL. F
35a.	8	8	10^{-8}		40	21	1.65	10^{-1}	10^{-5}	10^{-9}	REL. F
			10^{-12}		53	27	1.65	10^{-1}	10^{-7}	10^{-9}	REL. F
35b.[0]	9	9	10^{-8}		12	9	1.73	10^{-16}	10^{-15}	10^{-32}	X
			10^{-12}		13	10	1.73	10^{-16}	10^{-15}	10^{-32}	X
35c.	10	10	10^{-8}		25	12	1.81	10^{-1}	10^{-6}	10^{-3}	REL. F
			10^{-12}		34	17	1.81	10^{-1}	10^{-7}	10^{-3}	REL. F
36a.[0]	4	4	10^{-8}		(4000)	(3985)	27.9	10^{-7}	10^{-7}	10^{-13}	F LIM.
			10^{-12}		(4000)	(3985)	27.9	10^{-7}	10^{-7}	10^{-13}	F LIM.
36b.[0]	9	9	10^{-8}		(5310)	(5292)	30.8	10^{-7}	10^{-7}	10^{-14}	TIME
			10^{-12}		(5330)	(5312)	30.8	10^{-7}	10^{-7}	10^{-14}	TIME
36c.[0]	9	9	10^{-8}		29	28	1.73	10^{-17}	10^{-16}	10^{-33}	X
			10^{-12}		40	33	1.73	10^{-17}	10^{-17}	10^{-34}	X
36d.[0]	9	9	10^{-8}		(9000)	(8982)	39.2	10^{-7}	10^{-7}	10^{-14}	F LIM.
			10^{-12}		(9000)	(8982)	39.2	10^{-7}	10^{-7}	10^{-14}	F LIM.
37.	2	16	10^{-8}		15	14	8.85	10^{1}	10^{-1}	10^{-6}	REL. F
			10^{-12}		21	20	8.85	10^{1}	10^{-3}	10^{-6}	REL. F
38.	3	16	10^{-8}		18	16	26.1	10^{1}	10^{-2}	10^{-6}	REL. F
			10^{-12}		28	26	26.1	10^{1}	10^{-4}	10^{-6}	REL. F

Numerical Results for LMDER

	n	m	TOL	max. step	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
39a.	2	3	10^{-8}		5	4	10^{-6}	10^{-1}	10^{-7}	10^{-7}	REL. F
			10^{-12}		6	5	10^{-7}	10^{-1}	10^{-8}	10^{-7}	REL. F
39b.	2	3	10^{-8}		14	13	10^{-5}	10^{-1}	10^{-5}	10^{-7}	REL. F
			10^{-12}		21	20	10^{-6}	10^{-1}	10^{-7}	10^{-7}	REL. F
39c.	2	3	10^{-8}		18	10	10^{-5}	10^{-1}	10^{-5}	10^{-7}	REL. F
			10^{-12}		25	13	10^{-7}	10^{-1}	10^{-8}	10^{-7}	REL. F
39d.	2	3	10^{-8}		20	13	10^{-5}	10^{-1}	10^{-5}	10^{-7}	REL. F
			10^{-12}		28	18	10^{-7}	10^{-1}	10^{-7}	10^{-7}	REL. F
39e.	2	3	10^{-8}		28	19	10^{-5}	10^{-1}	10^{-4}	10^{-7}	REL. F
			10^{-12}		44	31	10^{-8}	10^{-1}	10^{-7}	10^{-7}	REL. F
39f.	2	3	10^{-8}		31	23	10^{-6}	10^{-1}	10^{-4}	10^{-7}	REL. F
			10^{-12}		44	33	10^{-8}	10^{-1}	10^{-6}	10^{-7}	REL. F
39g.	2	3	10^{-8}		39	29	10^{-8}	10^{-1}	10^{-4}	10^{-7}	REL. F
			10^{-12}		44	31	10^{-9}	10^{-1}	10^{-6}	10^{-7}	REL. F
40a.	3	4	10^{-8}		6	5	10^{-5}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		9	8	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
40b.	3	4	10^{-8}		14	8	10^{-4}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		17	10	10^{-7}	10^{0}	10^{-7}	10^{-7}	REL. F
40c.	3	4	10^{-8}		16	8	10^{-5}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		22	12	10^{-7}	10^{0}	10^{-7}	10^{-7}	REL. F
40d.	3	4	10^{-8}		26	17	10^{-5}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		40	27	10^{-7}	10^{0}	10^{-7}	10^{-7}	REL. F
40e.	3	4	10^{-8}		90	76	10^{-5}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		146	125	10^{-6}	10^{0}	10^{-5}	10^{-7}	REL. F
40f.	3	4	10^{-8}		180	158	10^{-5}	10^{0}	10^{-3}	10^{-7}	REL. F
			10^{-12}		272	241	10^{-7}	10^{0}	10^{-5}	10^{-7}	REL. F
40g.	3	4	10^{-8}		206	184	10^{-5}	10^{0}	10^{-2}	10^{-7}	REL. F
			10^{-12}		319	287	10^{-7}	10^{0}	10^{-5}	10^{-7}	REL. F
41a.	5	10	10^{-8}		4	3	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
			10^{-12}		4	3	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
41b.	5	10	10^{-8}		4	3	10^{-6}	10^{0}	10^{-7}	10^{-7}	REL. F
			10^{-12}		5	4	10^{-6}	10^{0}	10^{-9}	10^{-7}	REL. F
41c.	5	10	10^{-8}		6	5	10^{-6}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		8	7	10^{-6}	10^{0}	10^{-7}	10^{-7}	REL. F
41d.	5	10	10^{-8}		15	11	10^{-5}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		22	16	10^{-7}	10^{0}	10^{-6}	10^{-7}	REL. F
41e.	5	10	10^{-8}		29	18	10^{-5}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		38	24	10^{-7}	10^{0}	10^{-6}	10^{-7}	REL. F
41f.	5	10	10^{-8}		57	46	10^{-5}	10^{0}	10^{-3}	10^{-7}	REL. F
			10^{-12}		89	74	10^{-7}	10^{0}	10^{-5}	10^{-7}	REL. F
41g.	5	10	10^{-8}		84	71	10^{-5}	10^{0}	10^{-3}	10^{-7}	REL. F
			10^{-12}		144	123	10^{-7}	10^{0}	10^{-5}	10^{-7}	REL. F

	n	m	TOL	max. step	f evals.	iters./ J evals.	$\|r^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
42a.[0]	4	24	10^{-8}		18	15	60.8	10^{-13}	10^{-11}	10^{-26}	x
			10^{-12}		19	16	60.8	10^{-13}	10^{-11}	10^{-26}	x
42b.[0]	4	24	10^{-8}	0.001	48	42	61.3	10^{-13}	10^{-11}	10^{-25}	x
			10^{-12}	0.001	49	43	61.3	10^{-13}	10^{-11}	10^{-25}	x
42c.[0]	4	24	10^{-8}	0.1	20	17	60.3	10^{-13}	10^{-10}	10^{-26}	x
			10^{-12}	0.1	20	17	60.3	10^{-13}	10^{-10}	10^{-26}	x
42d.[0]	4	24	10^{-8}	0.1	15	14	60.3	10^{-13}	10^{-11}	10^{-26}	x
			10^{-12}	0.1	16	15	60.3	10^{-13}	10^{-11}	10^{-26}	x
43a.[0]	5	16	10^{-8}	0.1	14	11	54.0	10^{-14}	10^{-12}	10^{-27}	x
			10^{-12}	0.1	15	12	54.0	10^{-14}	10^{-12}	10^{-28}	x
43b.[0]	5	16	10^{-8}	0.1	18	15	54.0	10^{-14}	10^{-12}	10^{-28}	x
			10^{-12}	0.1	18	15	54.0	10^{-14}	10^{-12}	10^{-28}	x
43c.[0]	5	16	10^{-8}	0.1	11	10	54.0	10^{-14}	10^{-11}	10^{-27}	x
			10^{-12}	0.1	11	10	54.0	10^{-14}	10^{-11}	10^{-27}	x
43d.[0]	5	16	10^{-8}	0.1	22	18	54.0	10^{-14}	10^{-11}	10^{-27}	x
			10^{-12}	0.1	23	19	54.0	10^{-14}	10^{-12}	10^{-28}	x
43e.[0]	5	16	10^{-8}	0.1	12	11	54.0	10^{-14}	10^{-11}	10^{-27}	x
			10^{-12}	0.1	13	12	54.0	10^{-14}	10^{-12}	10^{-28}	x
43f.[0]	5	16	10^{-8}		12	9	54.0	10^{-14}	10^{-12}	10^{-27}	x
			10^{-12}		13	10	54.0	10^{-14}	10^{-12}	10^{-27}	x
44a.[0]	6	6	10^{-8}		37	30	4.06	10^{-15}	10^{-13}	10^{-30}	x
			10^{-12}		38	31	4.06	10^{-15}	10^{-13}	10^{-30}	x
44b.[0]	6	6	10^{-8}		5	4	3.52	10^{-15}	10^{-13}	10^{-29}	x
			10^{-12}		6	5	3.52	10^{-15}	10^{-13}	10^{-29}	x
44c.[0]	6	6	10^{-8}		108	98	20.6	10^{-14}	10^{-11}	10^{-29}	x
			10^{-12}		109	99	20.6	10^{-15}	10^{-11}	10^{-30}	x
44d.[0]	6	6	10^{-8}		98	88	15.3	10^{-15}	10^{-11}	10^{-29}	x
			10^{-12}		99	89	15.3	10^{-15}	10^{-11}	10^{-29}	x
44e.[0]	6	6	10^{-8}		82	71	9.27	10^{-14}	10^{-11}	10^{-28}	x
			10^{-12}		83	72	9.27	10^{-14}	10^{-11}	10^{-28}	x
45a.[0]	8	8	10^{-8}		47	35	4.06	10^{-16}	10^{-15}	10^{-29}	x
			10^{-12}		48	36	4.06	10^{-16}	10^{-15}	10^{-29}	x
45b.[0]	8	8	10^{-8}		5	4	3.56	10^{-15}	10^{-13}	10^{-29}	x
			10^{-12}		6	5	3.56	10^{-15}	10^{-13}	10^{-29}	x
45c.[0]	8	8	10^{-8}		164	148	20.6	10^{-14}	10^{-11}	10^{-28}	x
			10^{-12}		165	149	20.6	10^{-15}	10^{-11}	10^{-29}	x
45d.[0]	8	8	10^{-8}		144	133	15.3	10^{-15}	10^{-13}	10^{-30}	x
			10^{-12}		145	134	15.3	10^{-16}	10^{-13}	10^{-31}	x
45e.[0]	8	8	10^{-8}		130	119	9.31	10^{-14}	10^{-12}	10^{-29}	x
			10^{-12}		131	120	9.31	10^{-14}	10^{-12}	10^{-29}	x

5.7 Adaptive Special Quasi-Newton Method

(PORT/ACM DN2G/NL2SOL)

5.7.1 Software and Algorithm

The results were obtained using subroutine DN2G, a double precision version of the ACM algorithm NL2SOL available in the PORT Library [1984]. A subproblem of the form

$$\min_{p \in \Re^n} Q_k(p) \equiv g_k^T p + \frac{1}{2} p^T (J_k^T J_k + B_k) p$$

$$\text{subject to} \quad \|D_k p\|_2 \leq \delta_k$$

is solved at each iteration for the step p_k to the next iterate, where D_k is a diagonal scaling matrix. The method is adaptive, so that B_k is sometimes null and sometimes a scaled quasi-Newton approximation to the part of the Hessian involving the second derivatives of f.

5.7.2 Parameters

Parameters were kept at their default values with the following exceptions:

IV(MXFCAL)	-	$\min\{9999, 1000 * n\}$	function evaluation limit
IV(MXITER)	-	$\min\{9999, 1000 * n\}$	iteration limit
V(AFCTOL)	-	TOL*TOL (varied; see tables)	absolute function convergence tolerance
V(RFCTOL)	-	TOL (varied; see tables)	relative function convergence tolerance
V(SCTOL)	-	ϵ_M	singular convergence tolerance
V(XCTOL)	-	TOL (varied; see tables)	x convergence tolerance
V(XFTOL)	-	ϵ_M	false convergence tolerance
V(LMAX0)	-	usually 1.0 (default) †	initial trust-region diameter
V(LMAXS)	-	1.0 (default)	step bound for singular convergence test
V(TUNER1)	-	0.1 (default)	reduction test coefficient

† In some cases the default V(LMAX0) = 1.0 for the initial diameter of the trust-region was too large and overflow occurred during function evaluation. These cases are indicated in the table by giving the lower value of V(LMAX0) that was subsequently used to obtain the results in the column labeled "init. diam.".

See Dennis, Gay, and Welsch [1981a, 1981b], Gay [1983], and PORT [1984] for details concerning the parameters.

5.7.3 Convergence Criteria

The following quantities will be used in describing the convergence criteria:

objective function	:	$\mathcal{F}_k \equiv f_k^T f_k$
objective gradient	:	$g_k = \nabla\mathcal{F}_k \equiv 2J_k^T f_k$
current step	:	p_k, the minimizer of the subproblem
Newton step	:	$p_N \begin{cases} H_k^{-1}g & \text{if } H_k \text{ is positive definite;} \\ \textit{undefined} & \text{otherwise.} \end{cases}$
Newton reduction	:	$\rho_N = \begin{cases} -Q_k(p_N) & \text{if } H_k \text{ is positive definite;} \\ 0 & \text{otherwise.} \end{cases}$
predicted reduction	:	$\rho_P = -Q_k(p_k)$
actual reduction	:	$\rho_A = \mathcal{F}_k - \mathcal{F}(x_k + p_k)$
scaled distance	:	$\nu(x, y, D) = \dfrac{\max_{1\le i\le n} \{\lvert(D(x-y))^i\rvert\}}{\max_{1\le i\le n} \{\lvert(Dx)^i\rvert + \lvert(Dy)^i\rvert\}}$.†

† Here v_i denotes the ith component of the vector v. There is a provision for the user to replace the function ν; we used the default in all of the tests.

The convergence criteria used in DN2G are as follows:

- *Absolute function convergence* occurs at x_k if

$$|\mathcal{F}_k| < \mathtt{V(AFCTOL)}. \tag{5.7.1}$$

- *Relative function convergence* is intended to approximate the condition

$$\mathcal{F}_k - \mathcal{F}(x_*) \le \mathtt{V(RFCTOL)}\,|\mathcal{F}_k|.$$

The test actually used is

$$\rho_N \le \mathtt{V(RFCTOL)}\,|\mathcal{F}_k|. \tag{5.7.2}$$

- *x convergence* is intended to approximate the condition

$$\nu(x_k, x^*, D_k) \le \mathtt{V(XCTOL)}.$$

The test actually used is

$$p_k = p_N \text{ and } \nu(x_k, x_k + p_k, D_k) \le \mathtt{V(XCTOL)}. \tag{5.7.3}$$

- *Singular convergence* is intended to approximate the condition

$$\mathcal{F}_k - \min\{\mathcal{F}(y) \mid \|D_k(y - x_k)\| \le \mathtt{V(LMAXS)}\} < \mathtt{V(SCTOL)}\,|\mathcal{F}_k|,$$

where D_k is the diagonal scaling matrix at the kth iterate — when none of the convergence criteria listed above hold. It is meant to indicate relative function convegence when the Hessian in the subproblem is singular.

The actual test is

$$\mathcal{F}_k - \min\{Q_k(y) \mid \|D_k(y - x_k)\| \le \text{V(LMAXS)}\} < \text{V(SCTOL)}\,|\mathcal{F}_k|. \qquad (5.7.4)$$

Under certain conditions, the test is repeated for a step of length **V(LMAXS)**.

• *False convergence* is returned if none of the other convergence criteria are satisfied and a trial step no larger than **V(XFCTOL)** is rejected. This usually indicates either an error in computing the objective gradient, or a discontinuity (in $\mathcal{F}$ or g) near the current iterate, or that one or more of the convergence tolerances (**V(RFCTOL)**, **V(XCTOL)**, and **V(AFCTOL)**) are too small relative the accuracy to which the objective is computed.

The test actually used is

$$\mathcal{F}_k - \mathcal{F}(x_k + p_k) \le \text{V(TUNER1)}\rho_r \quad \text{and} \quad \nu(x_k, x_k + p_k, D_k) \le \text{V(XFTOL)}, \qquad (5.7.5)$$

where the parameter **V(TUNER1)** is adjustable, although in these tests the default value 0.1 is used throughout.

Except for test (5.6.13), tests for convergence are performed only when

$$\rho_A \le 2\rho_r. \qquad (5.7.6)$$

See Dennis, Gay, and Welsch [1981a, 1981b], Gay [1983], and PORT [1984] for more discussion of the convergence criteria.

The following abbreviations are used in the tables to describe the conditions under which the algorithm terminates:

ABS. F	-	(5.7.1)
REL. F	-	(5.7.2) and (5.7.6)
X	-	(5.7.3) and (5.7.6)
X, F	-	(5.7.2) and (5.7.3) and (5.7.6)
SING.	-	(5.7.4) and (5.7.6)
FALSE	-	(5.7.5) and (5.7.6)
F LIM.	-	function evaluation limit reached
TIME	-	time limit exceeded
LOOP	-	subroutine appears to loop

The total number of Jacobian evaluations is either equal to the total number of iterations of the method, or it is one more than the number of iterations. The number in the column labeled "iters. / J evals." is followed by a "+" if an extra Jacobian evaluation was used in the computation.

Numerical Results for DN2G

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\Vert x^*\Vert_2$	$\Vert f^*\Vert_2$	$\Vert g^*\Vert_2$	est. err.	conv.
1.[0]	2	2	10^{-8}		14	11+	1.41	10^{-16}	10^{-15}	10^{-32}	ABS. F
			10^{-12}		14	11+	1.41	10^{-16}	10^{-15}	10^{-32}	ABS. F
2.[0]	2	2	10^{-8}		10	8+	11.4	10^{1}	10^{-4}	10^{1}	REL. F
			10^{-12}		12	10	11.4	10^{1}	10^{-6}	10^{1}	REL. F
3.[0]	2	2	10^{-8}		64	50+	9.11	10^{-9}	10^{-4}	10^{-19}	ABS. F
			10^{-12}		65	51+	9.11	10^{-17}	10^{-12}	10^{-33}	ABS. F
4.[0]	2	3	10^{-8}		40	33+	10^{6}	10^{-2}	10^{-2}	10^{-4}	X
			10^{-12}		53	46+	10^{6}	10^{-6}	10^{-6}	10^{-12}	X
5.[0]	2	3	10^{-8}		9	7+	3.04	10^{-11}	10^{-10}	10^{-21}	ABS. F
			10^{-12}		10	8+	3.04	0.00	0.00	0.00	ABS. F
6.	2	10	10^{-8}		14	10+	.365	10^{1}	10^{-2}	10^{-6}	REL. F
			10^{-12}		16	12+	.365	10^{1}	10^{-7}	10^{-6}	REL. F
7.[0]	3	3	10^{-8}		13	8+	1.00	10^{-11}	10^{-9}	10^{-21}	ABS. F
			10^{-12}		14	9+	1.00	10^{-23}	10^{-21}	10^{-45}	ABS. F
8.	3	15	10^{-8}		7	6+	2.60	10^{-1}	10^{-9}	10^{-8}	REL. F
			10^{-12}		8	7+	2.60	10^{-1}	10^{-11}	10^{-8}	REL. F
9.	3	15	10^{-8}		3	2+	1.08	10^{-4}	10^{-12}	10^{-14}	X
			10^{-12}		5	4	1.08	10^{-4}	10^{-16}	10^{-14}	X, REL. F
10.	3	16	10^{-8}		132	120+	10^{4}	10^{1}	10^{-3}	10^{-6}	X, REL. F
			10^{-12}		133	121	10^{4}	10^{1}	10^{-3}	10^{-6}	X, REL. F
11.[0]	3	10	10^{-8}		(3000)	(2953)	239.	10^{-2}	10^{18}	10^{-5}	F LIM
			10^{-12}		(3000)	(2953)	239.	10^{-2}	10^{18}	10^{-5}	F LIM.
12.[0]	3	10	10^{-8}		8	5+	10.1	10^{-10}	10^{-10}	10^{-19}	ABS. F
			10^{-12}		9	6+	10.1	10^{-16}	10^{-16}	10^{-32}	ABS. F
13.[0]	4	4	10^{-8}		19	16+	10^{-4}	10^{-8}	10^{-12}	10^{-17}	ABS. F
			10^{-12}		25	22+	10^{-6}	10^{-12}	10^{-17}	10^{-24}	ABS. F
14.[0]	4	6	10^{-8}		52	40+	2.00	10^{-9}	10^{-7}	10^{-17}	ABS. F
			10^{-12}		53	41+	2.00	0.00	0.00	0.00	ABS. F
15.	4	11	10^{-8}		11	9+	.328	10^{-2}	10^{-9}	10^{-9}	REL. F
			10^{-12}		12	10+	.328	10^{-2}	10^{-9}	10^{-9}	REL. F
16.	4	20	10^{-8}		21	14+	17.6	10^{2}	10^{-2}	10^{-8}	REL. F
			10^{-12}		22	15+	17.6	10^{2}	10^{-4}	10^{-8}	REL. F
17.	5	33	10^{-8}		26	20+	2.46	10^{-2}	10^{-8}	10^{-11}	REL. F
			10^{-12}		27	21+	2.46	10^{-2}	10^{-10}	10^{-11}	REL. F
18.[0]	6	13	10^{-8}		45	32+	12.3	10^{-9}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		46	33+	12.3	10^{-16}	10^{-16}	10^{-31}	ABS. F
19.	11	65	10^{-8}		20	13+	9.38	10^{-1}	10^{-7}	10^{-8}	REL. F
			10^{-12}		22	15+	9.38	10^{-1}	10^{-9}	10^{-8}	REL. F

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
20a.	6	31	10^{-8}		13	9+	2.44	10^{-2}	10^{-9}	10^{-10}	REL. F
			10^{-12}		13	9+	2.44	10^{-2}	10^{-9}	10^{-10}	REL. F
20b.	9	31	10^{-8}		12	9+	6.06	10^{-3}	10^{-11}	10^{-13}	REL. F
			10^{-12}		15	12	6.06	10^{-3}	10^{-14}	10^{-13}	X, REL. F
20c.	12	31	10^{-8}		14	11+	16.6	10^{-5}	10^{-13}	10^{-16}	REL. F
			10^{-12}		14	11+	16.6	10^{-5}	10^{-13}	10^{-16}	REL. F
20d.	20	31	10^{-8}		8	6+	1.11	10^{-8}	10^{-13}	10^{-16}	ABS. F
			10^{-12}		(471)	(137+)	1.18	10^{-8}	10^{-14}	10^{-16}	LOOP
21a.[0]	10	10	10^{-8}		27	17+	3.16	10^{-16}	10^{-14}	10^{-31}	ABS. F
			10^{-12}		27	17+	3.16	10^{-16}	10^{-14}	10^{-31}	ABS. F
21b.[0]	20	20	10^{-8}		16	12+	4.47	10^{-16}	10^{-14}	10^{-31}	ABS. F
			10^{-12}		16	12+	4.47	10^{-16}	10^{-14}	10^{-31}	ABS. F
22a.[0]	12	12	10^{-8}		19	16+	10^{-4}	10^{-8}	10^{-12}	10^{-17}	ABS. F
			10^{-12}		26	23+	10^{-6}	10^{-12}	10^{-18}	10^{-24}	ABS. F
22b.[0]	20	20	10^{-8}		20	17+	10^{-4}	10^{-8}	10^{-12}	10^{-17}	ABS. F
			10^{-12}		26	23+	10^{-6}	10^{-12}	10^{-18}	10^{-24}	ABS. F
23a.	4	5	10^{-8}		36	26+	.500	10^{-3}	10^{-10}	10^{-10}	REL. F
			10^{-12}		37	27+	.500	10^{-3}	10^{-10}	10^{-10}	REL. F
23b.	10	11	10^{-8}		61	46+	.500	10^{-2}	10^{-7}	10^{-11}	REL. F
			10^{-12}		68	50+	.500	10^{-2}	10^{-10}	10^{-11}	REL. F
24a.	4	8	10^{-8}		139	110+	.759	10^{-3}	10^{-7}	10^{-11}	REL. F
			10^{-12}		142	113+	.759	10^{-3}	10^{-11}	10^{-11}	REL. F
24b.	10	20	10^{-8}		129	101+	.598	10^{-2}	10^{-7}	10^{-9}	REL. F
			10^{-12}		138	108+	.598	10^{-2}	10^{-8}	10^{-9}	REL. F
25a.[0]	10	12	10^{-8}		15	10+	3.16	10^{-12}	10^{-11}	10^{-24}	ABS. F
			10^{-12}		16	11+	3.16	10^{-15}	10^{-14}	10^{-31}	X
25b.[0]	20	22	10^{-8}		19	12+	4.47	10^{-15}	10^{-13}	10^{-29}	X
			10^{-12}		19	12+	4.47	10^{-15}	10^{-13}	10^{-29}	ABS. F
26a.[0]	10	10	10^{-8}		11	7+	.306	10^{-9}	10^{-10}	10^{-19}	ABS. F
			10^{-12}		12	8+	.306	10^{-15}	10^{-15}	10^{-30}	ABS. F
26b.[0]	20	20	10^{-8}		39	25+	.228	10^{-3}	10^{-9}	10^{-6}	REL. F
			10^{-12}		42	27	.228	10^{-3}	10^{-10}	10^{-6}	REL. F
27a.[0]	10	10	10^{-8}		8	6+	3.18	10^{-10}	10^{-10}	10^{-21}	ABS. F
			10^{-12}		9	7+	3.18	10^{-15}	10^{-14}	10^{-29}	ABS. F
27b.[0]	20	20	10^{-8}		11	8+	4.47	10^{-8}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		12	9+	4.47	10^{-14}	10^{-13}	10^{-27}	ABS. F
28a.[0]	10	10	10^{-8}		4	3+	.412	10^{-15}	10^{-16}	10^{-31}	ABS. F
			10^{-12}		4	3+	.412	10^{-15}	10^{-16}	10^{-31}	ABS. F
28b.[0]	20	20	10^{-8}		3	2+	.571	10^{-8}	10^{-9}	10^{-16}	ABS. F
			10^{-12}		4	3+	.571	10^{-16}	10^{-16}	10^{-32}	ABS. F

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
29a.[0]	10	10	10^{-8}		4	3+	.412	10^{-14}	10^{-14}	10^{-29}	ABS. F
			10^{-12}		4	3+	.412	10^{-14}	10^{-14}	10^{-29}	ABS. F
29b.[0]	20	20	10^{-8}		4	3+	.571	10^{-14}	10^{-14}	10^{-28}	ABS. F
			10^{-12}		4	3+	.571	10^{-14}	10^{-14}	10^{-28}	ABS. F
30a.[0]	10	10	10^{-8}		8	5+	2.05	10^{-9}	10^{-9}	10^{-18}	ABS. F
			10^{-12}		9	6+	2.05	10^{-16}	10^{-15}	10^{-31}	ABS. F
30b.[0]	20	20	10^{-8}		8	5+	3.04	10^{-8}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		9	6+	3.04	10^{-15}	10^{-14}	10^{-30}	ABS. F
31a.[0]	10	10	10^{-8}		10	7+	1.80	10^{-12}	10^{-11}	10^{-23}	ABS. F
			10^{-12}		11	8+	1.80	10^{-16}	10^{-15}	10^{-31}	X
31b.[0]	20	20	10^{-8}		10	7+	2.66	10^{-10}	10^{-10}	10^{-21}	ABS. F
			10^{-12}		11	8+	2.66	10^{-15}	10^{-14}	10^{-30}	ABS. F
32.[L]	10	20	10^{-8}		5	2	3.16	10^{0}	10^{-14}	10^{-16}	X, REL. F
			10^{-12}		5	2	3.16	10^{0}	10^{-14}	10^{-16}	X, REL. F
33.[L]	10	20	10^{-8}		18	5	20.2	10^{0}	10^{-11}	10^{-6}	SING.
			10^{-12}		18	5	20.2	10^{0}	10^{-11}	10^{-6}	SING.
34.[L]	10	20	10^{-8}		13	5	7.24	10^{0}	10^{-10}	10^{-6}	SING.
			10^{-12}		13	5	7.24	10^{0}	10^{-10}	10^{-6}	SING.
35a.	8	8	10^{-8}		23	14+	1.65	10^{-1}	10^{-7}	10^{-9}	REL. F
			10^{-12}		24	15+	1.65	10^{-1}	10^{-8}	10^{-9}	REL. F
35b.[0]	9	9	10^{-8}		11	7+	1.73	10^{-12}	10^{-12}	10^{-24}	ABS. F
			10^{-12}		11	7+	1.73	10^{-12}	10^{-12}	10^{-24}	ABS. F
35c.	10	10	10^{-8}		17	12+	1.81	10^{-1}	10^{-7}	10^{-3}	REL. F
			10^{-12}		19	14+	1.81	10^{-1}	10^{-8}	10^{-3}	REL. F
36a.[0]	4	4	10^{-8}		(4000)	(3988)	27.7	10^{-7}	10^{-7}	10^{-13}	F LIM.
			10^{-12}		(4000)	(3988)	27.7	10^{-7}	10^{-7}	10^{-13}	F LIM.
36b.[0]	9	9	10^{-8}		(9000)	(8977)	35.5	10^{-7}	10^{-8}	10^{-14}	F LIM.
			10^{-12}		(9000)	(8977)	35.5	10^{-7}	10^{-8}	10^{-14}	F LIM.
36c.[0]	9	9	10^{-8}		16	15+	1.73	10^{-9}	10^{-8}	10^{-17}	ABS. F
			10^{-12}		22	21+	1.73	10^{-12}	10^{-12}	10^{-24}	ABS. F
36d.[0]	9	9	10^{-8}		(9000)	(8966)	38.4	10^{-7}	10^{-7}	10^{-14}	F LIM.
			10^{-12}		(9000)	(8966)	38.4	10^{-7}	10^{-7}	10^{-14}	F LIM.
37.	2	16	10^{-8}		10	8+	8.85	10^{1}	10^{-3}	10^{-6}	REL. F
			10^{-12}		11	9+	8.85	10^{1}	10^{-5}	10^{-6}	REL. F
38.	3	16	10^{-8}		10	8+	26.1	10^{1}	10^{-2}	10^{-6}	REL. F
			10^{-12}		12	10+	26.1	10^{1}	10^{-7}	10^{-6}	REL. F

Numerical Results for DN2G

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
39a.	2	3	10^{-8}		5	4+	10^{-6}	10^{-1}	10^{-10}	10^{-7}	REL. F
			10^{-12}		5	4+	10^{-7}	10^{-1}	10^{-10}	10^{-7}	REL. F
39b.	2	3	10^{-8}		6	5+	10^{-7}	10^{-1}	10^{-7}	10^{-7}	REL. F
			10^{-12}		7	6+	10^{-7}	10^{-1}	10^{-9}	10^{-7}	REL. F
39c.	2	3	10^{-8}		7	4+	10^{-7}	10^{-1}	10^{-7}	10^{-7}	REL. F
			10^{-12}		8	5+	10^{-7}	10^{-1}	10^{-10}	10^{-7}	REL. F
39d.	2	3	10^{-8}		7	5+	10^{-7}	10^{-1}	10^{-6}	10^{-7}	REL. F
			10^{-12}		8	6+	10^{-7}	10^{-1}	10^{-10}	10^{-7}	REL. F
39e.	2	3	10^{-8}		9	6+	10^{-8}	10^{-1}	10^{-7}	10^{-7}	REL. F
			10^{-12}		10	7+	10^{-8}	10^{-1}	10^{-10}	10^{-7}	REL. F
39f.	2	3	10^{-8}		14	10+	10^{-8}	10^{-1}	10^{-6}	10^{-7}	REL. F
			10^{-12}		15	11+	10^{-9}	10^{-1}	10^{-9}	10^{-7}	REL. F
39g.	2	3	10^{-8}		18	12+	10^{-7}	10^{-1}	10^{-4}	10^{-7}	REL. F
			10^{-12}		20	14+	10^{-9}	10^{-1}	10^{-6}	10^{-7}	REL. F
40a.	3	4	10^{-8}		7	6	10^{-6}	10^{0}	10^{-11}	10^{-7}	REL. F
			10^{-12}		7	6	10^{-6}	10^{0}	10^{-11}	10^{-7}	REL. F
40b.	3	4	10^{-8}		7	5+	10^{-6}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		11	9	10^{-6}	10^{0}	10^{-14}	10^{-7}	REL. F
40c.	3	4	10^{-8}		9	6+	10^{-7}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		10	7+	10^{-7}	10^{0}	10^{-9}	10^{-7}	REL. F
40d.	3	4	10^{-8}		9	7+	10^{-7}	10^{0}	10^{-8}	10^{-7}	REL. F
			10^{-12}		9	7+	10^{-7}	10^{0}	10^{-8}	10^{-7}	REL. F
40e.	3	4	10^{-8}		10	9+	10^{-7}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		11	10+	10^{-7}	10^{0}	10^{-6}	10^{-7}	REL. F
40f.	3	4	10^{-8}		13	10+	10^{-8}	10^{0}	10^{-6}	10^{-7}	REL. F
			10^{-12}		14	11+	10^{-8}	10^{0}	10^{-7}	10^{-7}	REL. F
40g.	3	4	10^{-8}		23	16+	10^{-8}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		25	18+	10^{-9}	10^{0}	10^{-7}	10^{-7}	REL. F
41a.	5	10	10^{-8}		4	3+	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
			10^{-12}		4	3+	10^{-6}	10^{0}	10^{-10}	10^{-7}	REL. F
41b.	5	10	10^{-8}		4	3+	10^{-6}	10^{0}	10^{-7}	10^{-7}	REL. F
			10^{-12}		5	4+	10^{-6}	10^{0}	10^{-11}	10^{-7}	REL. F
41c.	5	10	10^{-8}		6	5+	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
			10^{-12}		6	5+	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
41d.	5	10	10^{-8}		9	7+	10^{-6}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		11	9+	10^{-6}	10^{0}	10^{-8}	10^{-7}	REL. F
41e.	5	10	10^{-8}		17	13+	10^{-6}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		20	16+	10^{-7}	10^{0}	10^{-8}	10^{-7}	REL. F
41f.	5	10	10^{-8}		24	19+	10^{-6}	10^{0}	10^{-4}	10^{-7}	REL. F
			10^{-12}		27	22+	10^{-7}	10^{0}	10^{-7}	10^{-7}	REL. F
41g.	5	10	10^{-8}		29	22+	10^{-7}	10^{0}	10^{-5}	10^{-7}	REL. F
			10^{-12}		30	23+	10^{-8}	10^{0}	10^{-6}	10^{-7}	REL. F

Numerical Results for DN2G

	n	m	TOL	init. diam.	f evals.	iters./ J evals.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
42a.^{0}	4	24	10^{-8}	0.9	29	19+	60.8	10^{-13}	10^{-10}	10^{-25}	X
			10^{-12}	0.9	29	19+	60.8	10^{-13}	10^{-10}	10^{-25}	ABS. F
42b.^{0}	4	24	10^{-8}	0.001	74	48+	61.3	10^{-13}	10^{-11}	10^{-25}	X
			10^{-12}	0.001	74	48+	61.3	10^{-13}	10^{-11}	10^{-25}	ABS. F
42c.^{0}	4	24	10^{-8}	0.01	32	19+	60.3	10^{-13}	10^{-10}	10^{-26}	X
			10^{-12}	0.01	32	19+	60.3	10^{-13}	10^{-10}	10^{-26}	ABS. F
42d.^{0}	4	24	10^{-8}		23	18+	60.3	10^{-10}	10^{-7}	10^{-19}	ABS. F
			10^{-12}		24	19+	60.3	10^{-13}	10^{-10}	10^{-26}	X
43a.^{0}	5	16	10^{-8}		31	22+	54.0	10^{-12}	10^{-10}	10^{-24}	ABS. F
			10^{-12}		32	23+	54.0	10^{-14}	10^{-11}	10^{-28}	X
43b.^{0}	5	16	10^{-8}		20	13+	54.0	10^{-12}	10^{-10}	10^{-24}	ABS. F
			10^{-12}		20	13+	54.0	10^{-12}	10^{-10}	10^{-24}	ABS. F
43c.^{0}	5	16	10^{-8}		34	26+	53.6	10^{-1}	10^{-7}	10^{-2}	REL. F
			10^{-12}		41	33+	53.6	10^{-1}	10^{-10}	10^{-2}	REL. F
43d.^{0}	5	16	10^{-8}		17	11+	54.0	10^{-14}	10^{-11}	10^{-27}	X
			10^{-12}		17	11+	54.0	10^{-14}	10^{-11}	10^{-27}	ABS. F
43e.^{0}	5	16	10^{-8}		28	18+	54.0	10^{-11}	10^{-9}	10^{-22}	ABS. F
			10^{-12}		29	19+	54.0	10^{-14}	10^{-12}	10^{-27}	X
43f.^{0}	5	16	10^{-8}		20	15+	54.0	10^{-12}	10^{-10}	10^{-25}	ABS. F
			10^{-12}		20	15+	54.0	10^{-12}	10^{-10}	10^{-25}	ABS. F
44a.^{0}	6	6	10^{-8}		58	41+	4.06	10^{-10}	10^{-8}	10^{-20}	ABS. F
			10^{-12}		59	42+	4.06	10^{-15}	10^{-13}	10^{-30}	ABS. F
44b.^{0}	6	6	10^{-8}		7	6+	3.52	10^{-14}	10^{-12}	10^{-28}	X
			10^{-12}		7	6+	3.52	10^{-14}	10^{-12}	10^{-28}	ABS. F
44c.^{0}	6	6	10^{-8}		93	84+	20.6	10^{-10}	10^{-6}	10^{-19}	ABS. F
			10^{-12}		94	85+	20.6	10^{-15}	10^{-11}	10^{-29}	ABS. F
44d.^{0}	6	6	10^{-8}		97	81+	15.3	10^{-11}	10^{-8}	10^{-22}	ABS. F
			10^{-12}		98	82+	15.3	10^{-14}	10^{-11}	10^{-28}	X
44e.^{0}	6	6	10^{-8}		83	72+	9.27	10^{-14}	10^{-12}	10^{-29}	X
			10^{-12}		83	72+	9.27	10^{-14}	10^{-12}	10^{-29}	ABS. F
45a.^{0}	8	8	10^{-8}		65	45+	4.06	10^{-11}	10^{-9}	10^{-22}	ABS. F
			10^{-12}		66	46+	4.06	10^{-17}	10^{-13}	10^{-33}	ABS. F
45b.^{0}	8	8	10^{-8}		8	7+	3.56	10^{-13}	10^{-12}	10^{-26}	ABS. F
			10^{-12}		8	7+	3.56	10^{-13}	10^{-12}	10^{-26}	ABS. F
45c.^{0}	8	8	10^{-8}		129	123+	20.6	10^{-8}	10^{-4}	10^{-16}	ABS. F
			10^{-12}		130	124+	20.6	10^{-15}	10^{-10}	10^{-29}	ABS. F
45d.^{0}	8	8	10^{-8}		168	144+	15.3	10^{-14}	10^{-11}	10^{-29}	X
			10^{-12}		168	144+	15.3	10^{-14}	10^{-11}	10^{-29}	ABS. F
45e.^{0}	8	8	10^{-8}		173	165+	9.31	10^{-15}	10^{-12}	10^{-29}	X
			10^{-12}		173	165+	9.31	10^{-15}	10^{-12}	10^{-29}	ABS. F

5.8 Corrected Gauss-Newton Methods

(NPL/NAG LSQSDN and LSQFDQ)

5.8.1 Software and Algorithms

The results were obtained using subroutines LSQSDN and LSQFDQ implementing corrected Gauss-Newton methods from the National Physical Laboratory, which are available at Stanford Linear Accelerator Center. A subproblem of the form

$$\min_{p \in \Re^n} g_k^T p + \frac{1}{2} p^T (J_k^T J_k + B_k) p$$

$$\text{subject to} \quad J_k p \approx -f_k,$$

is solved for a search direction p_k, where $\approx$ is interpreted in a least-squares sense using the singular-value decomposition (see Chapters 3 and 4). Subroutine LSQSDN requires exact second derivatives for the term B_k that involves the second derivatives of the residuals, while LSQFDQ uses a quasi-Newton approximation. The linesearch algorithm used within the subroutines requires both function and gradient information (see Gill and Murray [1974], for details). These subroutines are similar to those available in the NAG Library [1984] for solving nonlinear least-squares problems :LSQSDN corresponds to NAG subroutine E04HEF and LSQFDQ to NAG subroutine E04GBF.

5.8.2 Parmeters

LSQSDN and LSQFDQ have the same set of input parameters as the corresponding software from the NAG Library [1984]. The values chosen are listed below.

MAXCAL	-	min $\{9999, 1000 * n\}$	function evaluation limit
XTOL	-	varied; see tables	accuracy in x
ETA	-	0.9	linesearch accuracy
STEPMX	-	usually 10^6 (default) †	maximum step for linesearch

† In some cases the default STEPMX $= 10^6$ was too large and overflow occurred during function evaluation in the linesearch. These cases are indicated in the tables by giving the lower value of STEPMX that was subsequently used to obtain the results.

See the NAG [1984] manual for details concerning the parameters.

5.8.3 Convergence Criteria

The following quantities will be used in describing the convergence criteria :

objective function	:	$\mathcal{F}_k = f_k^T f_k$
objective gradient	:	$g_k = \nabla \mathcal{F}_k = 2J_k^T f_k$
search direction	:	p_k, the minimizer of the subproblem
steplength	:	α_k, determined by the linesearch

An iterate is determined to be optimal by LSQSDN and LSQFDQ if

$$\mathcal{F}_k < \epsilon_M^2 \tag{5.8.1}$$

or

$$\|g_k\|_2 < \epsilon_M \|f_k\|_2 \tag{5.8.2}$$

or if the following three conditions hold :

$$\alpha_k \|p_k\|_2 < (\mathrm{XTOL} + \epsilon_M)(1 + \|x_k\|_2) \tag{5.8.3}$$

and

$$\mathcal{F}(x_{k-1}) - \mathcal{F}_k < (\mathrm{XTOL} + \epsilon_M)^2 (1 + |\mathcal{F}_k|) \tag{5.8.4}$$

and

$$\|g_k\|_2 < \epsilon_M^{1/3} (1 + |\mathcal{F}_k|). \tag{5.8.5}$$

Conditions (5.8.3) and (5.8.4) are meant to ensure that the sequence x_k has converged, while conditions (5.8.2) and (5.8.5) are intended to test whether the necessary condition that the gradient vanish at a minimum is approximately satisfied at x_k. Condition (5.8.2) allows the algorithm to accept a point as a local mimimum if a more restrictive test on the necessary condition than (5.8.5) is met, even if conditions (5.8.3) and (5.8.4) do not hold. For the zero-residual case, condition (5.8.5) specifies that the method may also terminate when $\|f_k\|_2$ is no larger than the relative machine precision. For a detailed discussion of convergence criteria similar to these, see Sections 8.2 and 8.5 of Gill, Murray, and Wright [1981]. In particular, Section 8.5.1.3 treats special considerations relevant to nonlinear least squares.

The following abbreviations are used in the tables to describe the conditions under which the algorithm terminates :

OPT	-	optimal point found
*	-	current point cannot be improved †
F LIM	-	function evalutaion limit reached
TIME	-	time limit exceeded

† A '*' corresponds to the situation in which the algorithm terminates due to failure in the linesearch to find an acceptable step at the current iteration.

Numerical Results for LSQFDQ

	n	m	XTOL	max. step	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
1.^{0}	2	2	10^{-8}		31	12	1.41	0.00	0.00	0.00	OPT
			10^{-12}		31	12	1.41	0.00	0.00	0.00	OPT
2.^{0}	2	2	10^{-8}		36	26	11.4	10^{1}	10^{-8}	10^{1}	*
			10^{-12}		36	26	11.4	10^{1}	10^{-8}	10^{1}	*
3.^{0}	2	2	10^{-8}		47	25	9.11	10^{-10}	10^{-5}	10^{-20}	*
			10^{-12}		47	25	9.11	10^{-10}	10^{-5}	10^{-20}	*
4.^{0}	2	3	10^{-8}		64	20	10^{6}	10^{-9}	10^{-3}	10^{-18}	*
			10^{-12}		64	20	10^{6}	10^{-9}	10^{-3}	10^{-18}	*
5.^{0}	2	3	10^{-8}		14	9	3.04	10^{-14}	10^{-13}	10^{-28}	*
			10^{-12}		14	9	3.04	10^{-14}	10^{-13}	10^{-28}	*
6.	2	10	10^{-8}		54	28	.365	10^{1}	10^{-5}	10^{-6}	*
			10^{-12}		54	28	.365	10^{1}	10^{-5}	10^{-6}	*
7.^{0}	3	3	10^{-8}		20	13	1.00	10^{-11}	10^{-10}	10^{-23}	*
			10^{-12}		20	13	1.00	10^{-11}	10^{-10}	10^{-23}	*
8.	3	15	10^{-8}		13	12	2.60	10^{-1}	10^{-10}	10^{-8}	OPT.
			10^{-12}		13	12	2.60	10^{-1}	10^{-10}	10^{-8}	OPT.
9.	3	15	10^{-8}		3	2	1.08	10^{-4}	10^{-12}	10^{-14}	OPT.
			10^{-12}		3	2	1.08	10^{-4}	10^{-12}	10^{-14}	OPT.
10.	3	16	10^{-8}		18	10	10^{4}	10^{1}	10^{3}	10^{-6}	*
			10^{-12}		18	10	10^{4}	10^{1}	10^{3}	10^{-6}	*
11.^{0}	3	10	10^{-8}		69	31	60.8	10^{-2}	10^{-9}	10^{-4}	*
			10^{-12}		69	31	60.8	10^{-2}	10^{-9}	10^{-4}	*
12.^{0}	3	10	10^{-8}		12	8	10.1	10^{-10}	10^{-10}	10^{-19}	*
			10^{-12}		12	8	10.1	10^{-10}	10^{-10}	10^{-19}	*
13.^{0}	4	4	10^{-8}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
			10^{-12}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
14.^{0}	4	6	10^{-8}		81	58	2.00	10^{-9}	10^{-7}	10^{-17}	*
			10^{-12}		81	58	2.00	10^{-9}	10^{-7}	10^{-17}	*
15.	4	11	10^{-8}		30	22	.328	10^{-2}	10^{-9}	10^{-9}	OPT
			10^{-12}		30	22	.328	10^{-2}	10^{-9}	10^{-9}	OPT.
16.	4	20	10^{-8}		62	39	17.6	10^{2}	10^{-4}	10^{-8}	*
			10^{-12}		62	39	17.6	10^{2}	10^{-4}	10^{-8}	*
17.	5	33	10^{-8}		19	11	2.46	10^{-2}	10^{-7}	10^{-11}	*
			10^{-12}		19	11	2.46	10^{-2}	10^{-7}	10^{-11}	*
18.^{0}	6	13	10^{-8}	10^{2}	(6004)	(4108)	756.	10^{-1}	10^{-3}	10^{-1}	F LIM
			10^{-12}	10^{2}	(6004)	(4108)	756.	10^{-1}	10^{-3}	10^{-1}	F LIM.
19.	11	65	10^{-8}		33	22	9.38	10^{-1}	10^{-8}	10^{-8}	*
			10^{-12}		33	22	9.38	10^{-1}	10^{-8}	10^{-8}	*

Numerical Results for LSQFDQ

	n	m	XTOL	max. step	f, J evals.	iters.	$\|\|x^*\|\|_2$	$\|\|f^*\|\|_2$	$\|\|g^*\|\|_2$	est. err.	conv.
20a.	6	31	10^{-8}		32	28	2.44	10^{-2}	10^{-8}	10^{-10}	*
			10^{-12}		32	28	2.44	10^{-2}	10^{-8}	10^{-10}	*
20b.	9	31	10^{-8}		6	5	6.06	10^{-3}	10^{-11}	10^{-13}	OPT.
			10^{-12}		6	5	6.06	10^{-3}	10^{-11}	10^{-13}	OPT.
20c.	12	31	10^{-8}		6	5	16.6	10^{-5}	10^{-14}	10^{-16}	OPT.
			10^{-12}		6	5	16.6	10^{-5}	10^{-14}	10^{-16}	OPT.
20d.	20	31	10^{-8}		18	11	247.	10^{-10}	10^{-12}	10^{-24}	*
			10^{-12}		18	11	247.	10^{-10}	10^{-12}	10^{-24}	*
21a.^{0}	10	10	10^{-8}		31	12	3.16	0.00	0.00	0.00	OPT.
			10^{-12}		31	12	3.16	0.00	0.00	0.00	OPT.
21b.^{0}	20	20	10^{-8}		31	12	4.47	0.00	0.00	0.00	OPT.
			10^{-12}		31	12	4.47	0.00	0.00	0.00	OPT.
22a.^{0}	12	12	10^{-8}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
			10^{-12}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
22b.^{0}	20	20	10^{-8}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
			10^{-12}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
23a.	4	5	10^{-8}		80	52	.500	10^{-3}	10^{-10}	10^{-10}	OPT.
			10^{-12}		80	52	.500	10^{-3}	10^{-10}	10^{-10}	OPT.
23b.	10	11	10^{-8}		143	78	.500	10^{-2}	10^{-9}	10^{-11}	OPT.
			10^{-12}		143	78	.500	10^{-2}	10^{-9}	10^{-11}	OPT.
24a.	4	8	10^{-8}		411	370	.759	10^{-3}	10^{-10}	10^{-11}	OPT.
			10^{-12}		411	370	.759	10^{-3}	10^{-10}	10^{-11}	OPT.
24b.	10	20	10^{-8}	10^2	566	501	.598	10^{-2}	10^{-9}	10^{-9}	OPT.
			10^{-12}	10^2	566	501	.598	10^{-2}	10^{-9}	10^{-9}	OPT.
25a.^{0}	10	12	10^{-8}		16	12	3.16	10^{-8}	10^{-6}	10^{-15}	*
			10^{-12}		16	12	3.16	10^{-8}	10^{-6}	10^{-15}	*
25b.^{0}	20	22	10^{-8}		18	14	4.47	10^{-8}	10^{-7}	10^{-17}	*
			10^{-12}		18	14	4.47	10^{-8}	10^{-7}	10^{-17}	*
26a.^{0}	10	10	10^{-8}		22	11	.306	10^{-11}	10^{-11}	10^{-22}	*
			10^{-12}		22	11	.306	10^{-11}	10^{-11}	10^{-22}	*
26b.^{0}	20	20	10^{-8}		18	9	.208	10^{-10}	10^{-10}	10^{-19}	OPT
			10^{-12}		21	12	.208	10^{-10}	10^{-10}	10^{-19}	*
27a.^{0}	10	10	10^{-8}	10^2	26	13	3.18	10^{-13}	10^{-13}	10^{-26}	*
			10^{-12}	10^2	26	13	3.18	10^{-13}	10^{-13}	10^{-26}	*
27b.^{0}	20	20	10^{-8}	10.0	31	17	4.47	10^{-10}	10^{-10}	10^{-20}	*
			10^{-12}	10.0	31	17	4.47	10^{-10}	10^{-10}	10^{-20}	*
28a.^{0}	10	10	10^{-8}		4	3	.412	10^{-15}	10^{-16}	10^{-31}	OPT.
			10^{-12}		4	3	.412	10^{-15}	10^{-16}	10^{-31}	OPT.
28b.^{0}	20	20	10^{-8}		4	3	.571	10^{-16}	10^{-16}	10^{-32}	OPT
			10^{-12}		4	3	.571	10^{-16}	10^{-16}	10^{-32}	OPT.

Numerical Results for LSQFDQ

	n	m	XTOL	max. step	f, J evals.	iters.	$\lVert x^* \rVert_2$	$\lVert f^* \rVert_2$	$\lVert g^* \rVert_2$	est. err.	conv.
29a.[0]	10	10	10^{-8}		10	6	.412	10^{-14}	10^{-14}	10^{-29}	*
			10^{-12}		10	6	.412	10^{-14}	10^{-14}	10^{-29}	*
29b.[0]	20	20	10^{-8}		10	6	.571	10^{-14}	10^{-14}	10^{-28}	*
			10^{-12}		10	6	.571	10^{-14}	10^{-14}	10^{-28}	*
30a.[0]	10	10	10^{-8}		11	7	2.05	10^{-9}	10^{-9}	10^{-18}	*
			10^{-12}		11	7	2.05	10^{-9}	10^{-9}	10^{-18}	*
30b.[0]	20	20	10^{-8}		11	7	3.04	10^{-9}	10^{-9}	10^{-18}	*
			10^{-12}		11	7	3.04	10^{-9}	10^{-9}	10^{-18}	*
31a.[0]	10	10	10^{-8}		12	8	1.80	10^{-8}	10^{-7}	10^{-16}	*
			10^{-12}		12	8	1.80	10^{-8}	10^{-7}	10^{-16}	*
31b.[0]	20	20	10^{-8}		12	8	2.66	10^{-8}	10^{-7}	10^{-16}	*
			10^{-12}		12	8	2.66	10^{-8}	10^{-7}	10^{-16}	*
32.[L]	10	20	10^{-8}		2	1	3.16	10^{0}	10^{-14}	10^{-16}	OPT
			10^{-12}		2	1	3.16	10^{0}	10^{-14}	10^{-16}	OPT
33.[L]	10	20	10^{-8}		2	1	1.46	10^{0}	10^{-9}	10^{-6}	OPT.
			10^{-12}		2	1	1.46	10^{0}	10^{-9}	10^{-6}	OPT.
34.[L]	10	20	10^{-8}		2	1	1.78	10^{0}	10^{-9}	10^{-6}	OPT.
			10^{-12}		2	1	1.78	10^{0}	10^{-9}	10^{-6}	OPT.
35a.	8	8	10^{-8}	10.0	87	35	1.65	10^{-1}	10^{-8}	10^{-9}	*
			10^{-12}	10.0	87	35	1.65	10^{-1}	10^{-8}	10^{-9}	*
35b.[0]	9	9	10^{-8}		34	14	1.73	10^{-10}	10^{-10}	10^{-21}	*
			10^{-12}		34	14	1.73	10^{-10}	10^{-10}	10^{-21}	*
35c.	10	10	10^{-8}	10.0	73	33	1.81	10^{-1}	10^{-9}	10^{-3}	*
			10^{-12}	10.0	73	33	1.81	10^{-1}	10^{-9}	10^{-3}	*
36a.[0]	4	4	10^{-8}		44	25	50.0	10^{-15}	10^{-14}	10^{-31}	*
			10^{-12}		44	25	50.0	10^{-15}	10^{-14}	10^{-31}	*
36b.[0]	9	9	10^{-8}		44	25	50.0	10^{-15}	10^{-14}	10^{-31}	*
			10^{-12}		44	25	50.0	10^{-15}	10^{-14}	10^{-31}	*
36c.[0]	9	9	10^{-8}		28	27	1.73	10^{-16}	10^{-16}	10^{-32}	OPT
			10^{-12}		28	27	1.73	10^{-16}	10^{-16}	10^{-32}	OPT
36d.[0]	9	9	10^{-8}		424	100	232.	10^{-10}	10^{-9}	10^{-20}	*
			10^{-12}		424	100	232.	10^{-10}	10^{-9}	10^{-20}	*
37.	2	16	10^{-8}		25	23	8.85	10^{1}	10^{-4}	10^{-6}	*
			10^{-12}		25	23	8.85	10^{1}	10^{-4}	10^{-6}	*
38.	3	16	10^{-8}		30	26	26.1	10^{1}	10^{-6}	10^{-6}	*
			10^{-12}		30	26	26.1	10^{1}	10^{-6}	10^{-6}	*

Numerical Results for LSQFDQ

	n	m	XTOL	max. step	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
39a.	2	3	10^{-8}		17	16	10^{-6}	10^{-1}	10^{-9}	10^{-7}	OPT
			10^{-12}		17	16	10^{-6}	10^{-1}	10^{-9}	10^{-7}	OPT
39b.	2	3	10^{-8}		24	23	10^{-7}	10^{-1}	10^{-9}	10^{-7}	OPT.
			10^{-12}		24	23	10^{-7}	10^{-1}	10^{-9}	10^{-7}	OPT.
39c.	2	3	10^{-8}		22	18	10^{-7}	10^{-1}	10^{-8}	10^{-7}	OPT
			10^{-12}		23	19	10^{-7}	10^{-1}	10^{-9}	10^{-7}	OPT.
39d.	2	3	10^{-8}		31	26	10^{-7}	10^{-1}	10^{-8}	10^{-7}	OPT
			10^{-12}		32	27	10^{-7}	10^{-1}	10^{-9}	10^{-7}	OPT
39e.	2	3	10^{-8}		32	26	10^{-8}	10^{-1}	10^{-8}	10^{-7}	*
			10^{-12}		32	26	10^{-8}	10^{-1}	10^{-8}	10^{-7}	*
39f.	2	3	10^{-8}		43	27	10^{-9}	10^{-1}	10^{-8}	10^{-7}	*
			10^{-12}		43	27	10^{-9}	10^{-1}	10^{-8}	10^{-7}	*
39g.	2	3	10^{-8}		49	28	10^{-9}	10^{-1}	10^{-6}	10^{-7}	*
			10^{-12}		49	28	10^{-9}	10^{-1}	10^{-6}	10^{-7}	*
40a.	3	4	10^{-8}		18	17	10^{-6}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-12}		18	17	10^{-6}	10^{0}	10^{-8}	10^{-7}	OPT.
40b.	3	4	10^{-8}		19	17	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT
			10^{-12}		19	17	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
40c.	3	4	10^{-8}		27	22	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-12}		27	22	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
40d.	3	4	10^{-8}		33	26	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-12}		34	27	10^{-7}	10^{0}	10^{-9}	10^{-7}	OPT.
40e.	3	4	10^{-8}		70	39	10^{-7}	10^{0}	10^{-8}	10^{-7}	OPT.
			10^{-12}		72	40	10^{-7}	10^{0}	10^{-8}	10^{-7}	*
40f.	3	4	10^{-8}		92	43	10^{-8}	10^{0}	10^{-8}	10^{-7}	*
			10^{-12}		92	43	10^{-8}	10^{0}	10^{-8}	10^{-7}	*
40g.	3	4	10^{-8}	10^{3}	123	53	10^{-9}	10^{0}	10^{-6}	10^{-7}	*
			10^{-12}	10^{3}	123	53	10^{-9}	10^{0}	10^{-6}	10^{-7}	*
41a.	5	10	10^{-8}		8	7	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
			10^{-12}		8	7	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
41b.	5	10	10^{-8}		18	13	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT
			10^{-12}		18	13	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
41c.	5	10	10^{-8}		21	18	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT
			10^{-12}		21	18	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
41d.	5	10	10^{-8}		38	28	10^{-6}	10^{0}	10^{-8}	10^{-7}	*
			10^{-12}		38	28	10^{-6}	10^{0}	10^{-8}	10^{-7}	*
41e.	5	10	10^{-8}		47	37	10^{-7}	10^{0}	10^{-8}	10^{-7}	*
			10^{-12}		47	37	10^{-7}	10^{0}	10^{-8}	10^{-7}	*
41f.	5	10	10^{-8}		54	45	10^{-7}	10^{0}	10^{-8}	10^{-7}	*
			10^{-12}		54	45	10^{-7}	10^{0}	10^{-8}	10^{-7}	*
41g.	5	10	10^{-8}		62	52	10^{-8}	10^{0}	10^{-7}	10^{-7}	*
			10^{-12}		62	52	10^{-8}	10^{0}	10^{-7}	10^{-7}	*

Numerical Results for LSQFDQ

	n	m	XTOL	max. step	f, J evals.	iters.	$\|\|r^*\|\|_2$	$\|\|f^*\|\|_2$	$\|\|g^*\|\|_2$	est. err.	conv.
42a.^{0}	4	24	10^{-8}	10.0	73	49	63.5	10^{-7}	10^{-4}	10^{-13}	*
			10^{-12}	10.0	73	49	63.5	10^{-7}	10^{-4}	10^{-13}	*
42b.^{0}	4	24	10^{-8}	2.0	94	71	61.9	10^{-10}	10^{-7}	10^{-19}	*
			10^{-12}	2.0	94	71	61.9	10^{-10}	10^{-7}	10^{-19}	*
42c.^{0}	4	24	10^{-8}	5.0	46	30	60.3	10^{-5}	10^{-3}	10^{-11}	*
			10^{-12}	5.0	46	30	60.3	10^{-5}	10^{-3}	10^{-11}	*
42d.^{0}	4	24	10^{-8}	5.0	27	19	60.3	10^{-5}	10^{-3}	10^{-10}	*
			10^{-12}	5.0	27	19	60.3	10^{-5}	10^{-3}	10^{-10}	*
43a.^{0}	5	16	10^{-8}	10.0	33	18	54.0	10^{-9}	10^{-6}	10^{-17}	*
			10^{-12}	10.0	33	18	54.0	10^{-9}	10^{-6}	10^{-17}	*
43b.^{0}	5	16	10^{-8}	10.0	45	21	54.0	10^{-8}	10^{-6}	10^{-16}	*
			10^{-12}	10.0	45	21	54.0	10^{-8}	10^{-6}	10^{-16}	*
43c.^{0}	5	16	10^{-8}	10.0	33	18	54.0	10^{-8}	10^{-6}	10^{-16}	*
			10^{-12}	10.0	33	18	54.0	10^{-8}	10^{-6}	10^{-16}	*
43d.^{0}	5	16	10^{-8}	10.0	38	20	54.0	10^{-9}	10^{-7}	10^{-18}	*
			10^{-12}	10.0	38	20	54.0	10^{-9}	10^{-7}	10^{-18}	*
43e.^{0}	5	16	10^{-8}	10.0	27	14	54.0	10^{-10}	10^{-8}	10^{-20}	*
			10^{-12}	10.0	27	14	54.0	10^{-10}	10^{-8}	10^{-20r}	*
43f.^{0}	5	16	10^{-8}		31	19	54.0	10^{-8}	10^{-6}	10^{-16}	*
			10^{-12}		31	19	54.0	10^{-8}	10^{-5}	10^{-16}	*
44a.^{0}	6	6	10^{-8}		97	28	4.06	10^{-12}	10^{-10}	10^{-23}	*
			10^{-12}		97	28	4.06	10^{-12}	10^{-10}	10^{-23}	*
44b.^{0}	6	6	10^{-8}		10	6	3.52	10^{-8}	10^{-6}	10^{-15}	*
			10^{-12}		10	6	3.52	10^{-8}	10^{-6}	10^{-15}	*
44c.^{0}	6	6	10^{-8}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
			10^{-12}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
44d.^{0}	6	6	10^{-8}		40	17	15.3	10^{-9}	10^{-5}	10^{-17}	*
			10^{-12}		40	17	15.3	10^{-9}	10^{-5}	10^{-17}	*
44e.^{0}	6	6	10^{-8}		47	20	9.27	10^{-13}	10^{-10}	10^{-25}	*
			10^{-12}		47	20	9.27	10^{-13}	10^{-10}	10^{-25}	*
45a.^{0}	8	8	10^{-8}		97	28	4.06	10^{-12}	10^{-10}	10^{-23}	*
			10^{-12}		97	28	4.06	10^{-12}	10^{-10}	10^{-23}	*
45b.^{0}	8	8	10^{-8}		12	7	3.56	10^{-8}	10^{-6}	10^{-15}	*
			10^{-12}		12	7	3.56	10^{-8}	10^{-6}	10^{-15}	*
45c.^{0}	8	8	10^{-8}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
			10^{-12}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
45d.^{0}	8	8	10^{-8}		42	18	15.3	10^{-9}	10^{-5}	10^{-17}	*
			10^{-12}		42	18	15.3	10^{-9}	10^{-5}	10^{-17}	*
45e.^{0}	8	8	10^{-8}		49	21	9.31	10^{-13}	10^{-10}	10^{-26}	*
			10^{-12}		49	21	9.31	10^{-13}	10^{-10}	10^{-26}	*

Numerical Results for LSQSDN

	n	m	XTOL	max. step	f, J evals.	iters.	$\|\|x^*\|\|_2$	$\|\|f^*\|\|_2$	$\|\|g^*\|\|_2$	est. err.	conv.
1.°	2	2	10^{-8}		31	12	1.41	0.00	0.00	0.00	OPT
			10^{-12}		31	12	1.41	0.00	0.00	0.00	OPT.
2.°	2	2	10^{-8}		18	8	11.4	10^{1}	10^{-9}	10^{1}	OPT
			10^{-12}		18	8	11.4	10^{1}	10^{-9}	10^{1}	OPT
3.°	2	2	10^{-8}		47	25	9.11	10^{-10}	10^{-5}	10^{-20}	*
			10^{-12}		47	25	9.11	10^{-10}	10^{-5}	10^{-20}	*
4.°	2	3	10^{-8}		53	21	10^{6}	10^{-9}	10^{-3}	10^{-18}	*
			10^{-12}		53	21	10^{6}	10^{-9}	10^{-3}	10^{-18}	*
5.°	2	3	10^{-8}		10	7	3.04	10^{-14}	10^{-13}	10^{-28}	OPT
			10^{-12}		10	7	3.04	10^{-14}	10^{-13}	10^{-28}	OPT
6.	2	10	10^{-8}	5.0	36	10	.365	10^{1}	10^{-5}	10^{-6}	*
			10^{-12}	5.0	36	10	.365	10^{1}	10^{-5}	10^{-6}	*
7.°	3	3	10^{-8}		14	10	1.00	10^{-11}	10^{-10}	10^{-23}	OPT.
			10^{-12}		14	10	1.00	10^{-11}	10^{-10}	10^{-23}	OPT.
8.	3	15	10^{-8}		6	5	2.60	10^{-1}	10^{-9}	10^{-8}	OPT
			10^{-12}		6	5	2.60	10^{-1}	10^{-9}	10^{-8}	OPT.
9.	3	15	10^{-8}		3	2	1.08	10^{-4}	10^{-12}	10^{-14}	OPT
			10^{-12}		3	2	1.08	10^{-4}	10^{-12}	10^{-14}	OPT.
10.	3	16	10^{-8}		17	10	10^{4}	10^{1}	10^{2}	10^{-6}	*
			10^{-12}		17	10	10^{4}	10^{1}	10^{2}	10^{-6}	*
11.°	3	10	10^{-8}	10.0	30	15	49.0	10^{-2}	10^{-9}	10^{-3}	*
			10^{-12}	10.0	30	15	49.0	10^{-2}	10^{-9}	10^{-3}	*
12.°	3	10	10^{-8}		8	6	10.1	10^{-10}	10^{-10}	10^{-19}	OPT
			10^{-12}		12	8	10.1	10^{-10}	10^{-10}	10^{-19}	*
13.°	4	4	10^{-8}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
			10^{-12}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
14.°	4	6	10^{-8}		87	42	2.00	10^{-8}	10^{-7}	10^{-16}	OPT
			10^{-12}		93	45	2.00	10^{-8}	10^{-7}	10^{-16}	*
15.	4	11	10^{-8}		16	8	.328	10^{-2}	10^{-14}	10^{-9}	OPT
			10^{-12}		16	8	.328	10^{-2}	10^{-14}	10^{-9}	OPT.
16.	4	20	10^{-8}		45	22	17.6	10^{2}	10^{-6}	10^{-8}	*
			10^{-12}		45	22	17.6	10^{2}	10^{-6}	10^{-8}	*
17.	5	33	10^{-8}		14	9	2.46	10^{-2}	10^{-9}	10^{-11}	OPT
			10^{-12}		18	11	2.46	10^{-2}	10^{-9}	10^{-11}	*
18.°	6	13	10^{-8}	10.0	243	69	12.3	10^{-12}	10^{-11}	10^{-23}	OPT.
			10^{-12}	10.0	247	71	12.3	10^{-12}	10^{-11}	10^{-23}	*
19.	11	65	10^{-8}		19	10	9.38	10^{-1}	10^{-9}	10^{-8}	OPT.
			10^{-12}		19	10	9.38	10^{-1}	10^{-9}	10^{-8}	OPT.

Numerical Results for LSQSDN

	n	m	XTOL	max. step	f, J evals.	iters.	$\Vert x^*\Vert_2$	$\Vert f^*\Vert_2$	$\Vert g^*\Vert_2$	est. err.	conv.
20a.	6	31	10^{-8}		9	7	2.44	10^{-2}	10^{-8}	10^{-10}	*
			10^{-12}		9	7	2.44	10^{-2}	10^{-8}	10^{-10}	*
20b.	9	31	10^{-8}		6	5	6.06	10^{-3}	10^{-11}	10^{-13}	OPT.
			10^{-12}		6	5	6.06	10^{-3}	10^{-11}	10^{-13}	OPT.
20c.	12	31	10^{-8}		6	5	16.6	10^{-5}	10^{-14}	10^{-16}	OPT.
			10^{-12}		6	5	16.6	10^{-5}	10^{-14}	10^{-16}	OPT.
20d.	20	31	10^{-8}		10	8	247.	10^{-10}	10^{-11}	10^{-24}	OPT.
			10^{-12}		12	9	247.	10^{-10}	10^{-11}	10^{-24}	OPT.
21a.[0]	10	10	10^{-8}		31	12	3.16	0.00	0.00	0.00	OPT.
			10^{-12}		31	12	3.16	0.00	0.00	0.00	OPT.
21b.[0]	20	20	10^{-8}		31	12	4.47	0.00	0.00	0.00	OPT.
			10^{-12}		31	12	4.47	0.00	0.00	0.00	OPT.
22a.[0]	12	12	10^{-8}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
			10^{-12}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
22b.[0]	20	20	10^{-8}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
			10^{-12}		18	17	10^{-5}	10^{-9}	10^{-13}	10^{-18}	OPT.
23a.	4	5	10^{-8}		58	32	.500	10^{-3}	10^{-11}	10^{-10}	OPT.
			10^{-12}		58	32	.500	10^{-3}	10^{-11}	10^{-10}	OPT.
23b.	10	11	10^{-8}		124	64	.500	10^{-2}	10^{-13}	10^{-11}	OPT.
			10^{-12}		124	64	.500	10^{-2}	10^{-13}	10^{-11}	OPT.
24a.	4	8	10^{-8}		228	136	.759	10^{-3}	10^{-11}	10^{-11}	OPT.
			10^{-12}		228	136	.759	10^{-3}	10^{-11}	10^{-11}	OPT.
24b.	10	20	10^{-8}		150	88	.598	10^{-2}	10^{-11}	10^{-9}	OPT.
			10^{-12}		150	88	.598	10^{-2}	10^{-11}	10^{-9}	OPT.
25a.[0]	10	12	10^{-8}		12	10	3.16	10^{-8}	10^{-6}	10^{-15}	OPT.
			10^{-12}		17	12	3.16	10^{-8}	10^{-6}	10^{-15}	*
25b.[0]	20	22	10^{-8}		14	12	4.47	10^{-8}	10^{-7}	10^{-17}	OPT.
			10^{-12}		19	14	4.47	10^{-8}	10^{-7}	10^{-17}	*
26a.[0]	10	10	10^{-8}		18	9	.306	10^{-11}	10^{-11}	10^{-22}	OPT.
			10^{-12}		22	11	.306	10^{-11}	10^{-11}	10^{-22}	*
26b.[0]	20	20	10^{-8}		18	9	.208	10^{-10}	10^{-10}	10^{-19}	OPT.
			10^{-12}		26	12	.208	10^{-10}	10^{-10}	10^{-19}	*
27a.[0]	10	10	10^{-8}	10^2	22	7	3.18	10^{-10}	10^{-9}	10^{-20}	OPT.
			10^{-12}	10^2	28	10	3.18	10^{-10}	10^{-9}	10^{-20}	*
27b.[0]	20	20	10^{-8}	10^2	21	7	4.47	10^{-9}	10^{-9}	10^{-19}	OPT.
			10^{-12}	10^2	27	10	4.47	10^{-9}	10^{-9}	10^{-19}	*
28a.[0]	10	10	10^{-8}		4	3	.412	10^{-15}	10^{-16}	10^{-31}	OPT.
			10^{-12}		4	3	.412	10^{-15}	10^{-16}	10^{-31}	OPT.
28b.[0]	20	20	10^{-8}		4	3	.571	10^{-16}	10^{-16}	10^{-32}	OPT.
			10^{-12}		4	3	.571	10^{-16}	10^{-16}	10^{-32}	OPT.

	n	m	XTOL	max. step	f, J evals.	iters.	$\|x^*\|_2$	$\|f^*\|_2$	$\|g^*\|_2$	est. err.	conv.
29a.[0]	10	10	10^{-8}		6	4	.412	10^{-14}	10^{-14}	10^{-29}	OPT.
			10^{-12}		6	4	.412	10^{-14}	10^{-14}	10^{-29}	OPT.
29b.[0]	20	20	10^{-8}		6	4	.571	10^{-14}	10^{-14}	10^{-28}	OPT.
			10^{-12}		6	4	.571	10^{-14}	10^{-14}	10^{-28}	OPT.
30a.[0]	10	10	10^{-8}		7	5	2.05	10^{-9}	10^{-9}	10^{-18}	OPT.
			10^{-12}		11	7	2.05	10^{-9}	10^{-9}	10^{-18}	*
30b.[0]	20	20	10^{-8}		7	5	3.04	10^{-9}	10^{-9}	10^{-18}	OPT.
			10^{-12}		11	7	3.04	10^{-9}	10^{-9}	10^{-18}	*
31a.[0]	10	10	10^{-8}		8	6	1.80	10^{-8}	10^{-7}	10^{-16}	OPT.
			10^{-12}		12	8	1.80	10^{-8}	10^{-7}	10^{-16}	*
31b.[0]	20	20	10^{-8}		8	6	2.66	10^{-8}	10^{-7}	10^{-16}	OPT.
			10^{-12}		12	8	2.66	10^{-8}	10^{-7}	10^{-16}	*
32.[L]	10	20	10^{-8}		2	1	3.16	10^{0}	10^{-14}	10^{-16}	OPT.
			10^{-12}		2	1	3.16	10^{0}	10^{-14}	10^{-16}	OPT.
33.[L]	10	20	10^{-8}		2	1	1.46	10^{0}	10^{-9}	10^{-6}	OPT.
			10^{-12}		2	1	1.46	10^{0}	10^{-9}	10^{-6}	OPT.
34.[L]	10	20	10^{-8}		2	1	1.78	10^{0}	10^{-9}	10^{-6}	OPT.
			10^{-12}		2	1	1.78	10^{0}	10^{-9}	10^{-6}	OPT.
35a.	8	8	10^{-8}		74	19	1.65	10^{-1}	10^{-7}	10^{-9}	*
			10^{-12}		74	19	1.65	10^{-1}	10^{-7}	10^{-9}	*
35b.[0]	9	9	10^{-8}		30	12	1.73	10^{-11}	10^{-10}	10^{-22}	OPT.
			10^{-12}		34	14	1.73	10^{-11}	10^{-10}	10^{-22}	*
35c.	10	10	10^{-8}		43	10	1.81	10^{-1}	10^{-10}	10^{-3}	OPT.
			10^{-12}		43	10	1.81	10^{-1}	10^{-10}	10^{-3}	OPT.
36a.[0]	4	4	10^{-8}		38	22	50.0	10^{-15}	10^{-14}	10^{-31}	OPT.
			10^{-12}		38	22	50.0	10^{-15}	10^{-14}	10^{-31}	OPT.
36b.[0]	9	9	10^{-8}		38	22	50.0	10^{-15}	10^{-14}	10^{-31}	OPT.
			10^{-12}		38	22	50.0	10^{-15}	10^{-14}	10^{-31}	OPT.
36c.[0]	9	9	10^{-8}		28	27	1.73	10^{-16}	10^{-16}	10^{-32}	OPT.
			10^{-12}		28	27	1.73	10^{-16}	10^{-16}	10^{-32}	OPT.
36d.[0]	9	9	10^{-8}		380	84	233.	10^{-10}	10^{-9}	10^{-20}	OPT.
			10^{-12}		380	84	233.	10^{-10}	10^{-9}	10^{-20}	OPT.
37.	2	16	10^{-8}		9	8	8.85	10^{1}	10^{-10}	10^{-6}	OPT.
			10^{-12}		9	8	8.85	10^{1}	10^{-10}	10^{-6}	OPT.
38.	3	16	10^{-8}		10	8	26.1	10^{1}	10^{-9}	10^{-6}	OPT.
			10^{-12}		10	8	26.1	10^{1}	10^{-9}	10^{-6}	OPT.

Numerical Results for LSQSDN

	n	m	XTOL	max. step	f, J evals.	iters.	$\lVert x^*\rVert_2$	$\lVert f^*\rVert_2$	$\lVert g^*\rVert_2$	est. err.	conv.
39a.	2	3	10^{-8}		4	3	10^{-6}	10^{-1}	10^{-9}	10^{-7}	OPT.
			10^{-12}		4	3	10^{-6}	10^{-1}	10^{-9}	10^{-7}	OPT.
39b.	2	3	10^{-8}		6	5	10^{-7}	10^{-1}	10^{-15}	10^{-7}	OPT.
			10^{-12}		6	5	10^{-7}	10^{-1}	10^{-15}	10^{-7}	OPT.
39c.	2	3	10^{-8}		9	5	10^{-7}	10^{-1}	10^{-10}	10^{-7}	OPT
			10^{-12}		9	5	10^{-7}	10^{-1}	10^{-10}	10^{-7}	OPT
39d.	2	3	10^{-8}		12	7	10^{-7}	10^{-1}	10^{-10}	10^{-7}	OPT
			10^{-12}		12	7	10^{-7}	10^{-1}	10^{-10}	10^{-7}	OPT
39e.	2	3	10^{-8}		12	7	10^{-8}	10^{-1}	10^{-10}	10^{-7}	OPT
			10^{-12}		12	7	10^{-8}	10^{-1}	10^{-10}	10^{-7}	OPT.
39f.	2	3	10^{-8}		25	10	10^{-9}	10^{-1}	10^{-13}	10^{-7}	OPT
			10^{-12}		25	10	10^{-9}	10^{-1}	10^{-13}	10^{-7}	OPT
39g.	2	3	10^{-8}		39	15	10^{-10}	10^{-1}	10^{-7}	10^{-7}	*
			10^{-12}		39	15	10^{-10}	10^{-1}	10^{-7}	10^{-7}	*
40a.	3	4	10^{-8}		5	4	10^{-6}	10^{0}	10^{-18}	10^{-7}	OPT.
			10^{-12}		5	4	10^{-6}	10^{0}	10^{-15}	10^{-7}	OPT.
40b.	3	4	10^{-8}		6	4	10^{-6}	10^{0}	10^{-10}	10^{-7}	OPT.
			10^{-12}		6	4	10^{-6}	10^{0}	10^{-10}	10^{-7}	OPT.
40c.	3	4	10^{-8}		11	6	10^{-7}	10^{0}	10^{-14}	10^{-7}	OPT.
			10^{-12}		11	6	10^{-7}	10^{0}	10^{-14}	10^{-7}	OPT.
40d.	3	4	10^{-8}		13	7	10^{-7}	10^{0}	10^{-16}	10^{-7}	OPT.
			10^{-12}		13	7	10^{-7}	10^{0}	10^{-16}	10^{-7}	OPT.
40e.	3	4	10^{-8}		45	16	10^{-7}	10^{0}	10^{-13}	10^{-7}	OPT.
			10^{-12}		45	16	10^{-7}	10^{0}	10^{-13}	10^{-7}	OPT
40f.	3	4	10^{-8}		49	18	10^{-8}	10^{0}	10^{-15}	10^{-7}	OPT.
			10^{-12}		49	18	10^{-8}	10^{0}	10^{-15}	10^{-7}	OPT.
40g.	3	4	10^{-8}		85	24	10^{-9}	10^{0}	10^{-10}	10^{-7}	OPT.
			10^{-12}		85	24	10^{-9}	10^{0}	10^{-10}	10^{-7}	OPT.
41a.	5	10	10^{-8}		4	3	10^{-6}	10^{0}	10^{-12}	10^{-7}	OPT.
			10^{-12}		4	3	10^{-6}	10^{0}	10^{-12}	10^{-7}	OPT.
41b.	5	10	10^{-8}		4	3	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
			10^{-12}		4	3	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
41c.	5	10	10^{-8}		5	4	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT
			10^{-12}		5	4	10^{-6}	10^{0}	10^{-9}	10^{-7}	OPT.
41d.	5	10	10^{-8}		9	7	10^{-6}	10^{0}	10^{-12}	10^{-7}	OPT.
			10^{-12}		9	7	10^{-6}	10^{0}	10^{-12}	10^{-7}	OPT.
41e.	5	10	10^{-8}		14	10	10^{-7}	10^{0}	10^{-11}	10^{-7}	OPT
			10^{-12}		14	10	10^{-7}	10^{0}	10^{-11}	10^{-7}	OPT.
41f.	5	10	10^{-8}		16	12	10^{-7}	10^{0}	10^{-10}	10^{-7}	OPT.
			10^{-12}		16	12	10^{-7}	10^{0}	10^{-10}	10^{-7}	OPT.
41g.	5	10	10^{-8}		21	15	10^{-8}	10^{0}	10^{-12}	10^{-7}	OPT.
			10^{-12}		21	15	10^{-8}	10^{0}	10^{-12}	10^{-7}	OPT.

Numerical Results for LSQSDN

	n	m	XTOL	max. step	f, J evals.	iters.	$\lVert x^*\rVert_2$	$\lVert f^*\rVert_2$	$\lVert g^*\rVert_2$	est. err.	conv.
42a.[0]	4	24	10^{-8}	5.0	60	41	63.5	10^{-10}	10^{-7}	10^{-19}	OPT
			10^{-12}	5.0	64	43	63.5	10^{-10}	10^{-7}	10^{-19}	*
42b.[0]	4	24	10^{-8}	2.0	75	67	61.9	10^{-6}	10^{-3}	10^{-11}	*
			10^{-12}	2.0	75	67	61.9	10^{-6}	10^{-3}	10^{-11}	*
42c.[0]	4	24	10^{-8}	5.0	48	28	61.8	10^{-12}	10^{-9}	10^{-23}	OPT
			10^{-12}	5.0	48	28	61.8	10^{-12}	10^{-9}	10^{-23}	OPT
42d.[0]	4	24	10^{-8}	5.0	27	19	60.3	10^{-5}	10^{-3}	10^{-10}	*
			10^{-12}	5.0	27	19	60.3	10^{-5}	10^{-3}	10^{-10}	*
43a.[0]	5	16	10^{-8}	10.0	25	14	54.0	10^{-9}	10^{-6}	10^{-17}	OPT
			10^{-12}	10.0	33	18	54.0	10^{-9}	10^{-6}	10^{-17}	*
43b.[0]	5	16	10^{-8}	10.0	37	17	54.0	10^{-8}	10^{-6}	10^{-16}	OPT
			10^{-12}	10.0	45	21	54.0	10^{-8}	10^{-6}	10^{-16}	*
43c.[0]	5	16	10^{-8}	10.0	25	14	54.0	10^{-8}	10^{-6}	10^{-16}	OPT.
			10^{-12}	10.0	33	18	54.0	10^{-8}	10^{-6}	10^{-16}	*
43d.[0]	5	16	10^{-8}	10.0	30	16	54.0	10^{-9}	10^{-7}	10^{-18}	OPT.
			10^{-12}	10.0	38	20	54.0	10^{-9}	10^{-7}	10^{-18}	*
43e.[0]	5	16	10^{-8}	10.0	19	10	54.0	10^{-10}	10^{-8}	10^{-20}	OPT
			10^{-12}	10.0	27	14	54.0	10^{-10}	10^{-8}	10^{-20}	*
43f.[0]	5	16	10^{-8}		23	15	54.0	10^{-8}	10^{-6}	10^{-16}	OPT.
			10^{-12}		31	19	54.0	10^{-8}	10^{-6}	10^{-16}	*
44a.[0]	6	6	10^{-8}		93	26	4.06	10^{-12}	10^{-10}	10^{-28}	OPT.
			10^{-12}		95	27	4.06	10^{-12}	10^{-10}	10^{-28}	OPT.
44b.[0]	6	6	10^{-8}		6	4	3.52	10^{-8}	10^{-6}	10^{-15}	OPT.
			10^{-12}		10	6	3.52	10^{-8}	10^{-6}	10^{-15}	*
44c.[0]	6	6	10^{-8}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
			10^{-12}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
44d.[0]	6	6	10^{-8}		40	17	15.3	10^{-9}	10^{-5}	10^{-17}	*
			10^{-12}		40	17	15.3	10^{-9}	10^{-5}	10^{-17}	*
44e.[0]	6	6	10^{-8}		43	18	9.27	10^{-13}	10^{-10}	10^{-25}	OPT
			10^{-12}		43	18	9.27	10^{-13}	10^{-10}	10^{-25}	OPT
45a.[0]	8	8	10^{-8}		93	26	4.06	10^{-12}	10^{-10}	10^{-23}	OPT
			10^{-12}		95	27	4.06	10^{-12}	10^{-10}	10^{-23}	OPT.
45b.[0]	8	8	10^{-8}		6	4	3.56	10^{-8}	10^{-6}	10^{-15}	OPT
			10^{-12}		12	7	3.56	10^{-8}	10^{-6}	10^{-15}	*
45c.[0]	8	8	10^{-8}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
			10^{-12}		47	19	20.6	10^{-7}	10^{-3}	10^{-14}	*
45d.[0]	8	8	10^{-8}		42	18	15.3	10^{-9}	10^{-5}	10^{-17}	*
			10^{-12}		42	18	15.3	10^{-9}	10^{-5}	10^{-17}	*
45e.[0]	8	8	10^{-8}		43	18	9.31	10^{-13}	10^{-10}	10^{-25}	OPT.
			10^{-12}		43	18	9.31	10^{-13}	10^{-10}	10^{-25}	OPT.

5.9 Test Problems

Superscripts on problem numbers have the following interpretation :

^{0} : zero-residual problem
$^{\iota}$: linear least-squares problem

Problems from Moré, Garbow, and Hillstrom [1981]

	n	m	
1.^{0}	2	2	**Rosenbrock**
2.^{0}	2	2	**Freudenstein and Roth**
3.^{0}	2	2	**Powell Badly Scaled**
4.^{0}	2	3	**Brown Badly Scaled**
5.^{0}	2	3	**Beale**
6.	2	10	**Jennrich and Sampson**
7.^{0}	3	3	**Helical Valley**
8.	3	15	**Bard**
9.	3	15	**Gaussian**
10.	3	16	**Meyer**
11.^{0}	3	10	**Gulf Research and Development†**
12.^{0}	3	10	**Box 3-Dimensional**
13.^{0}	4	4	**Powell Singular**
14.^{0}	4	6	**Wood**
15.	4	11	**Kowalik and Osborne**
16.	4	20	**Brown and Dennis**
17.	5	33	**Osborne 1**
18.^{0}	6	13	**Biggs EXP6‡**

† For the **Gulf Research and Development Function (# 11)**, the formula

$$\phi_i(x) = \exp\left[-\frac{|y_i\ mi\ x_2|^{x_3}}{x_1}\right] - t_i$$

given in Moré, Garbow, and Hillstrom [1981] for the residual functions is in error. The correct formula is

$$\phi_i(x) = \exp\left[-\frac{|y_i - x_2|^{x_3}}{x_1}\right] - t_i$$

(see Moré, Garbow, and Hillstrom [1978]).

‡ For the **Biggs EXP6 Function (# 18)**, the minmum value for the sum of squares is given in Moré, Garbow, and Hillstrom [1981] as $5.65565\ldots \times 10^{-3}$. It can be easily verified that the residuals vanish at several points (for example $(1, 10, 1, 5, 4, 3)$).

Problems from Moré, Garbow, and Hillstrom [1981] (continued)

	n	m	
19.	11	65	Osborne 2†
20a.	6	31	Watson
20b.	9	31	Watson
20c.	12	31	Watson
20d.	20	31	Watson
21a.[0]	10	10	Extended Rosenbrock
21b.[0]	20	20	Extended Rosenbrock
22a.[0]	12	12	Extended Powell Singular
22b.[0]	20	20	Extended Powell Singular
23a.	4	5	Penalty I
23b.	10	11	Penalty I
24a.	4	8	Penalty II
24b.	10	20	Penalty II
25a.[0]	10	12	Variably Dimensioned
25b.[0]	20	22	Variably Dimensioned
26a.[0]	10	10	Trigonometric
26b.[0]	20	20	Trigonometric
27a.[0]	10	10	Brown Almost Linear
27b.[0]	20	20	Brown Almost Linear
28a.[0]	10	10	Discrete Boundary Value
28b.[0]	20	20	Discrete Boundary Value
29a.[0]	10	10	Discrete Integral
29b.[0]	20	20	Discrete Integral
30a.[0]	10	10	Broyden Tridiagonal
30b.[0]	20	20	Broyden Tridiagonal
31a.[0]	10	10	Broyden Banded
31b.[0]	20	20	Broyden Banded
32.[L]	10	20	Linear — Full Rank
33.[L]	10	20	Linear — Rank 1
34.[L]	10	20	Linear — Rank 1 with Zero Columns and Rows
35a.	8	8	Chebyquad
35b.[0]	9	9	Chebyquad
35c.	10	10	Chebyquad

† For Osborne's Second Function (# 19), the value of $f(x^*)$ is given (to six figures) in Moré, Garbow, and Hillstrom [1981] as 4.01377×10^{-2}. The smallest value we were able to obtain was 4.01683×10^{-2}.

Matrix Square Root Problems†‡

	n	m	
36a.^{0}	4	4	**Matrix Square Root 1**
36b.^{0}	9	9	**Matrix Square Root 2**
36c.^{0}	9	9	**Matrix Square Root 3**
36d.^{0}	9	9	**Matrix Square Root 4**

† These test problems come from a private communication of S. Hammarling to P. E. Gill in 1983.

	MATRIX	SQUARE ROOT
36a.^{0}	$\begin{pmatrix} 10^{-4} & 1 \\ 0 & 10^{-4} \end{pmatrix}$	$\begin{pmatrix} 10^{-2} & 50 \\ 0 & 10^{-2} \end{pmatrix}$
36b.^{0}	$\begin{pmatrix} 10^{-4} & 1 & 0 \\ 0 & 10^{-4} & 0 \\ 0 & 0 & 10^{-4} \end{pmatrix}$	$\begin{pmatrix} 10^{-2} & 50 & 0 \\ 0 & 10^{-2} & 0 \\ 0 & 0 & 10^{-2} \end{pmatrix}$
36c.^{0}	$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
36d.^{0}	$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

‡ The identity matrix was used as the starting value in all instances. Note that the iteration should not be started with the zero matrix because it is a stationary point of the sum of squares.

Problems from Salane [1987]

	n	m	
37.	2	16	**Hanson 1**
38.	3	16	**Hanson 2**

Problems from McKeown [1975a] (also McKeown [1975b])

	n	m		μ
39a.	2	3	**McKeown 1**	0.001
39b.	2	3	**McKeown 1**	0.01
39c.	2	3	**McKeown 1**	0.1
39d.	2	3	**McKeown 1**	1.0
39e.	2	3	**McKeown 1**	10.0
39f.	2	3	**McKeown 1**	100.0
39g.	2	3	**McKeown 1**	1000.0
40a.†	3	4	**McKeown 2**	0.001
40b.†	3	4	**McKeown 2**	0.01
40c.†	3	4	**McKeown 2**	0.1
40d.†	3	4	**McKeown 2**	1.0
40e.†	3	4	**McKeown 2**	10.0
40f.†	3	4	**McKeown 2**	100.0
40g.†	3	4	**McKeown 2**	1000.0
41a.	5	10	**McKeown 3**	0.001
41b.	5	10	**McKeown 3**	0.01
41c.	5	10	**McKeown 3**	0.1
41d.	5	10	**McKeown 3**	1.0
41e.	5	10	**McKeown 3**	10.0
41f.	5	10	**McKeown 3**	100.0
41g.	5	10	**McKeown 3**	1000.0

† In the data defining this problem given in McKeown [1975a] and [1975b], the matrix

$$B = \begin{pmatrix} 2.95137 & 4.87407 & -2.0506 \\ 4.87407 & 9.39321 & -3.93181 \\ -2.0506 & -3.93189 & 2.64745 \end{pmatrix}$$

is in error (it should be symmetric). The value

$$B = \begin{pmatrix} 2.95137 & 4.87407 & -2.0506 \\ 4.87407 & 9.39321 & -3.93189 \\ -2.0506 & -3.93189 & 2.64745 \end{pmatrix},$$

which is correct to six decimal digits, was used in our formulation of the problem.

Problems from DeVilliers and Glasser [1981] (also Salane [1987])

	n	m		starting value
42a.^{0}	4	24	**DeVilliers and Glasser 1**	(1.0, 8.0, 4.0, 4.412)
42b.^{0}	4	24	**DeVilliers and Glasser 1**	(1.0, 8.0, 8.0, 1.0)
42c.^{0}	4	24	**DeVilliers and Glasser 1**	(1.0, 8.0, 1.0, 4.412)
42d.^{0}	4	24	**DeVilliers and Glasser 1**	(1.0, 8.0, 4.0, 1.0)
43a.^{0}	5	16	**DeVilliers and Glasser 2**	(45.0, 2.0, 2.5, 1.5, 0.9)
43b.^{0}	5	16	**DeVilliers and Glasser 2**	(42.0, 0.8, 1.4, 1.8, 1.0)
43c.^{0}	5	16	**DeVilliers and Glasser 2**	(45.0, 2.0, 2.1, 2.0, 0.9)
43d.^{0}	5	16	**DeVilliers and Glasser 2**	(45.0, 2.5, 1.7, 1.0, 1.0)
43e.^{0}	5	16	**DeVilliers and Glasser 2**	(35.0, 2.5, 1.7, 1.0, 1.0)
43f.^{0}	5	16	**DeVilliers and Glasser 2**	(42.0, 0.8, 1.8, 3.15, 1.0)

Problems from Dennis, Gay, and Vu [1985]

	n	m		starting value
44a.^{0}†	6	6	**Exp. 791129**	(.299, −0.273, −.474, .474, −.0892, .0882)‡
44b.^{0}†	6	6	**Exp. 791226**	(−.3, .3, −1.2, 2.69, 1.59, −1.5)
44c.^{0}†	6	6	**Exp. 0121a**	(−.041, .03, −2.565, 2.565, −.754, .754)‡
44d.^{0}†	6	6	**Exp. 0121b**	(−.056, .026, −2.991, 2.991, −.568, .568)
44e.^{0}†	6	6	**Exp. 0121c**	(−.074, .013, −3.632, 3.632, −.289, .289)
45a.^{0}	8	8	**Exp. 791129**	(.299, .186, −0.273, .0254, −0.474, −.0892, .0892)‡
45b.^{0}	8	8	**Exp. 791226**	(−.3, −.39, .3, −.344, −1.2, 2.69, 1.59, −1.5)
45c.^{0}	8	8	**Exp. 0121a**	(−.041, −.775, .03, −.047, −2.565, 2.565, −.754, .754)‡
45d.^{0}	8	8	**Exp. 0121b**	(−.056, −.753, .026, −.047, −2.991, 2.991, −.568, .568)
45e.^{0}	8	8	**Exp. 0121c**	(−.074, −.733, .013, −.034, −3.632, 3.632, −.289, .289)

† Variables x_2 and x_4 (b and d in Dennis, Gay, and Vu [1985]) are eliminated from the linear constraints in order to get the 6-variable formulation of the problem (see Dennis, Gay, and Vu [1985]).

‡ Specification of some starting values in Dennis, Gay, and Vu [1985] is incomplete. The correct values were obtained from D. M. Gay in 1986.

6. Bibliography

Al-Baali, M., and R. Fletcher, "Variational Methods for Non-Linear Least Squares", *Journal of the Operational Research Society*, Vol. 36 No. 5 (May 1985) 405–421.

Armijo, L., "Minimization of Functions having Lipschitz-Continuous First Partial Derivatives", *Pacific Journal of Mathematics* 16 (1966) 1–3.

Brodlie, K. W., "An Assessment of two Approaches to Variable Metric Methods", *Mathematical Programming* 12 (1977) 344–355.

Brown, K. M., and J. E. Dennis, "A New Algorithm for Nonlinear Least Squares Curve Fitting", in *Mathematical Software*, J. R. Rice (ed.), Academic Press (1971).

Broyden, C. G., J. E. Dennis, and J. J. Moré, "On the Local and Superlinear Convergence of Quasi-Newton Methods", *Journal of the Institute of Mathematics and its Applications* 12 (1973) 223–246.

Bulteau, J. P., and J. P. Vial, "A Restricted Trust Region Algorithm for Unconstrained Minimization", *Journal of Optimization Theory and Applications*, Vol. 47 No. 4 (December 1985) 413–435.

Bulteau, J. P., and J. P. Vial, "Curvilinear Path and Trust Region in Unconstrained Optimization: A Convergence Analysis", *Mathematical Programming Studies* 30 (1987) 82–101.

Byrd, R. H., R. B. Schnabel, and G. A. Shultz, "Approximate Solution of the Trust Region Problem by Minimization over Two-Dimensional Subspaces", *Mathematical Programming*, Vol. 40 No. 3 (March 1988).

Davidon, W. C., "Optimally Conditioned Optimization Algorithms without Line Searches", *Mathematical Programming* 9 (1975) 1–30.

Dennis, J. E., and J. J. Moré, "A Characterization of Superlinear Convergence and its Application to Quasi-Newton Methods", *Mathematics of Computation*, Vol. 28 No. 126 (April 1974) 549–560.

Dennis, J. E., and J. J. Moré, "Quasi-Newton Methods: Motivation and Theory", *SIAM Review*, Vol. 19 No. 1 (January 1977) 46–89.

Dennis, J. E., and H. H. W. Mei, "Two New Unconstrained Optimization Methods which use Function and Gradient Values", *Journal of Optimization Theory and Applications* 28 (1979) 453–482.

Dennis, J. E., and R. B. Schnabel, "Least Change Secant Updates for Quasi-Newton Methods", *SIAM Review*, Vol. 21 No. 4 (October 1979) 443–459.

Dennis, J. E., D. M. Gay, and R. E. Welsch, "An Adaptive Nonlinear Least-Squares Algorithm", ACM *Transactions on Mathematical Software*, Vol. 7 No. 3 (September 1981a) 348–368.

Dennis, J. E., D. M. Gay, and R. E. Welsch, "ALGORITHM 573 NL2SOL : An Adaptive Nonlinear Least-Squares Algorithm", ACM *Transactions on Mathematical Software*, Vol. 7 No. 3 (September 1981b) 369–383.

Dennis, J. E., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall (1983).

Dennis, J. E., D. M. Gay, and R. E. Welsch, "Remark on ALGORITHM 573", ACM *Transactions on Mathematical Software*, Vol. 9 No. 1 (March 1983) 139.

Dennis, J. E., D. M. Gay, and P. A. Vu, "A New Nonlinear Equations Test Problem", Technical Report 83-16, Mathematical Sciences Department, Rice University (June 1983 — revised May 1985).

Flachs, J., "On the Generation of Updates for Quasi-Newton Methods", *Journal of Optimization Theory and Applications*, Vol. 48 No. 3 (March 1986) 379–417.

Fletcher, R., "A Modified Marquardt Subroutine for Non-Linear Least Squares", Technical Report AERE- R 6799, Atomic Energy Research Establishment, Harwell U. K. (1971).

Fletcher, R., and T. L. Freeman, "A Modified Newton Method for Minimization", *Journal of Optimization Theory and Applications*, Vol. 23 No. 3 (November 1977) 357–372.

Fletcher, R., *Practical Methods of Optimization : Volume I , Unconstrained Optimization*, Wiley (1980).

Fraley, C., "Solution of Nonlinear Least-Squares Problems", Technical Report STAN-CS-87-1165 (Ph. D. Thesis), Department of Computer Science, Stanford University, and Report CLaSSiC-87-04, Center for Large Scale Scientific Computation, Stanford University (July 1987a).

Fraley, C., "Computational Behavior of Gauss-Newton Methods", Technical Report SOL 87-10, Systems Optimization Laboratory, Department of Operations Research, Stanford University, and Manuscript CLaSSiC-87-21, Center for Large Scale Scientific Computation, Stanford University (August 1987b).

Fraley, C., "Algorithms for Nonlinear Least-Squares Problems", Technical Report, Systems Optimization Laboratory, Department of Operations Research, Stanford University; Center for Large Scale Scientific Computation, Stanford University; Université de Genève, COMIN (April 1988).

Gay, D. M., "Computing Optimal Locally Constrained Steps", SIAM *Journal on Scientific and Statistical Computing*, Vol. 2 No. 2 (June 1981) 186–197.

Gay, D. M., "ALGORITHM 611 : Subroutines for Unconstrained Minimization Using a Model/Trust-Region Approach", *ACM Transactions on Mathematical Software*, Vol. 9 No. 4 (December 1983) 503–524.

Gill, P. E., and W. Murray, "Quasi-Newton Methods for Unconstrained Optimization", *Journal of the Institute of Mathematics and its Applications* 9 (1972) 91–108.

Gill, P. E., and W. Murray, "Newton-type Methods for Unconstrained and Linearly Constrained Optimization", *Mathematical Programming* 7 (1974a) 311–350.

Gill, P. E., and W. Murray, "Safeguarded Steplength Algorithms for Optimization using Descent Methods", Technical Report NAC 37, National Physical Laboratory, England (1974b).

Gill, P. E., and W. Murray, "Nonlinear Least Squares and Nonlinearly Constrained Optimization", in *Numerical Analysis — Proceedings Dundee 1975, Lecture Notes in Mathematics* **506**, Springer-Verlag (1976) 134–147.

Gill, P. E., and W. Murray, "Algorithms for the Solution of the Nonlinear Least-Squares Problem", SIAM *Journal on Numerical Analysis*, Vol. 15 No. 5 (October 1978) 977–992.

Gill, P. E., W. Murray, M. A. Saunders, and M. H. Wright, "Two Steplength Algorithms for Numerical Optimization", Technical Report SOL 79-25, Systems Optimization Laboratory, Department of Operations Research, Stanford University (1979).

Gill, P. E., W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press (1981)

Gill, P. E., W. Murray, M. A. Saunders, and M. H. Wright, "Model Building and Practical Aspects of Nonlinear Programming", in NATO ASI Series, Vol. **F15**, *Computational Mathematical Programming*, K. Schittkowski (ed.), Springer-Verlag (1985) 209–247.

Gill, P. E., S. J. Hammarling, W. Murray, M. A. Saunders, and M. H. Wright, "User's Guide for LSSOL (Version 1.0): A Fortran Package for Constrained Linear Least-Squares and Convex Quadratic Programming", Technical Report SOL 86-1, Systems Optimization Laboratory, Department of Operations Research, Stanford University (January 1986a).

Gill, P. E., W. Murray, M. A. Saunders, and M. H. Wright, "Users Guide for NPSOL (Version 4.0): A Fortran Package for Nonlinear Programming", Technical Report SOL 86-2, Systems Optimization Laboratory, Department of Operations Research, Stanford University (January 1986b), revised Version 4.2 (1987).

Goldfarb, D., "Curvilinear Path Steplength Algorithms for Minimization which use Directions of Negative Curvature", *Mathematical Programming* 18 (1980) 31–40.

Goldfeld, S., R. Quandt, and H. Trotter, "Maximization by Quadratic Hill-Climbing", *Econometrica* 34 (1966) 541–551.

Goldstein, A. A., "On Steepest Descent", SIAM *Journal on Control* 3 (1965) 147–151.

Goldstein, A. A., *Constructive Real Analysis*, Harper and Row (1967)

Goldstein, A. A., and J. F. Price, "An Effective Algorithm for Minimization", *Numerische Mathematik* 10 (1967) 184–189.

Golub, G. H., and C. L. Van Loan, *Matrix Computations*, Johns-Hopkins (1983)

Hanson, R. J., and F. T. Krogh, "Testing Nonlinear Least Squares System Solvers Using Performance Profiling", Technical Report, Jet Propulsion Laboratory, Pasadena, CA U. S. A. (April 1987).

Hebden, M. D., "An Algorithm for Minimization using Exact Second Derivatives", Technical Report T. P. 515, Atomic Energy Research Establishment, Harwell U. K. (1973).

Hiebert, K. L., "A Comparison of Nonlinear Least Squares Software", Technical Report SAND79-0483, Sandia National Laboratories, Albuquerque, NM U. S. A. (1979).

Hiebert, K. L., "An Evaluation of Mathematical Software that Solves Nonlinear Least Squares Problems", ACM *Transactions on Mathematical Software*, Vol. 7 No. 1 (March 1981) 1–16.

Higham, N. J., "Computing the Polar Decomposition — with Applications", SIAM *Journal on Scientific and Statistical Computing*, Vol. 7 No. 4 (October 1986) 1160–1174.

Holt, J. N., and R. Fletcher, "An Algorithm for Constrained Nonlinear Least-Squares", *Journal of the Institute of Mathematics and its Applications* 23 (1979) 449–463.

Lawson, C. L., and R. J. Hanson, *Solving Least Squares Problems*, Prentice Hall (1974)

Levenberg, K., "A Method for the Solution of Certain Nonlinear Problems in Least Squares", *Quarterly of Applied Mathematics* 2 (1944) 164–168.

Marquardt, D. W., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", *Journal of the Society for Industrial and Applied Mathematics* (SIAM), Vol. 11 No. 2 (June 1963) 431–441.

McCormick, G. P., "A Modification of Armijo's Step-Size Rule for Negative Curvature", *Mathematical Programming* 13 (1977) 111–115.

Moré, J. J., B. S. Garbow, and K. E. Hillstrom, "Testing Unconstrained Optimization Software", Technical Report TM-324, Applied Mathematics Division, Argonne National Laboratory (July 1978).

Moré, J. J., B. S. Garbow, and K. E. Hillstrom, "User Guide for MINPACK-1", Technical Report ANL-80-74, Applied Mathematics Division, Argonne National Laboratory (August 1980).

Moré, J. J., and D. C. Sorensen, "On the Use of Directions of Negative Curvature in a Modified Newton Method", *Mathematical Programming* 16 (1979) 1–20.

Moré, J. J., B. S. Garbow, and K. E. Hillstrom, "Testing Unconstrained Optimization Software", *ACM Transactions on Mathematical Software*, Vol. 7 No. 1 (March 1981) 17–41.

Moré, J. J., "Recent Developments in Algorithms and Software for Trust Region Methods", in *Mathematical Programming—The State of the Art—Bonn 1982*, A. Bachem, M. Grötschel, and B. Korte (eds.), Springer-Verlag (1983) 258–287.

Moré, J. J., and D. C. Sorensen, "Computing a Trust Region Step", SIAM *Journal on Scientific and Statistical Computing*, Vol. 4 No. 3 (September 1983) 553–572.

Moré, J. J., and D. C. Sorensen, "Newton's Method", in *Studies in Numerical Analysis, MAA Studies in Mathematics*, Vol. 24, G. H. Golub (ed.), The Mathematical Association of America (1984).

Morrison, D. D., "Methods for Nonlinear Least Squares Problems and Convergence Proofs, Tracking Programs and Orbital Determinations", *Procedings of the Jet Propulsion Laboratory Seminar* (1960) 1–9.

NAG Fortran Library Manual, Vol. 3, Numerical Algorithms Group, Oxford U. K., and Downers Grove, IL U. S. A. (January 1984).

Nazareth, L., "An Alternative Variational Principle for Variable Metric Updating", *Mathematical Programming* 30 (1984) 99–104.

Oren, S. S., " Self-Scaling Variable Metric Algorithms Without Linesearch for Unconstrained Minimization", *Mathematics of Computation*, Vol. 27 No. 124 (October 1973).

Oren, S. S., and E. Spedicato, "Optimal Conditioning of Self-Scaling Variable Mettric Algorithms", *Mathematical Programming* 10 (1796) 70–90.

Oren, S. S., "Perspectives on Self-Scaling Variable Metric Algorithms", *Journal of Optimization Theory and Applications*, Vol. 37 No. 2 (June 1982) 137–148.

Ortega, J. M., and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press (1970).

Osborne, M. R., "Some Aspects of Non-linear Least Squares Calculations", in *Numerical Methods for Nonlinear Optimization*, F. A. Lootsma (ed.), Academic Press (1972) 171–189.

PORT Mathematical Subroutine Library Manual, A. T. & T. Bell Laboratories, Inc., Murray Hill, NJ U. S. A. (May 1984).

Powell, M. J. D., "A Hybrid Method for Nonlinear Equations", in *Numerical Methods for Nonlinear Algebraic Equations*, P. Rabinowitz (ed.), Gordon and Breach (1970a) 87–114.

Powell, M. J. D., "A New Algorithm for Unconstrained Optimization", in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (eds.) (1970b) 31–65.

Powell, M. J. D., "Convergence Properties of a Class of Minimization Algorithms", in *Nonlinear Programming 2*, O. L. Mangasarian, R. R. Meyer, and S. M. Robinson (eds.), Academic Press (1975) 1–27.

Reinsch, C. H., "Smoothing by Spline Functions II", *Numerische Mathematik* 16 (1971) 451–454.

Salane, D. E., "A Continuation Approach for Solving Large Residual Nonlinear Least Squares Problems", SIAM *Journal on Scientific and Statistical Computing*, Vol. 8 No. 4 (July 1987) 655–671.

Schnabel, R. B., "Optimal Conditioning in the Convex Class of Rank Two Updates", *Mathematical Programming* 15 (1978) 247–260.

Shanno, D. F., and P. C. Kettler, "Optimal Conditioning of Quasi-Newton Methods", *Mathematics of Computation* 24 (1970) 657–667.

Shanno, D. F., and K. H. Phua, "Matrix Conditioning and Nonlinear Optimization", *Mathematical Programming* 14 (1978a) 145–160.

Shanno, D. F., and K. H. Phua, "Numerical Comparison of Several Variable Metric Algorithms", *Journal of Optimization Theory and Applications* 25 (1978b) 507–518.

Shultz, G. A., R. B. Schnabel, and R. H. Byrd, "A Family of Trust-Region-Based Algorithms for Unconstrained Minimization with Strong Global Convergence Properties", SIAM *Journal on Numerical Analysis*, Vol. 22 No. 1 (February 1985) 47–67.

Sorensen, D. C., "Newton's Method with a Model Trust Region Modification", SIAM *Journal on Numerical Analysis*, Vol. 19 No. 2 (April 1982) 409–426.

Sorensen, D. C., "Trust Region Methods for Unconstrained Minimization", in *Nonlinear Optimization 1981*, M. J. D. Powell (ed.), Academic Press (1982) 29–38.

Spedicato, E., "A Variable Metric Method for Function Minimization Derived from Invariancy to Nonlinear Scaling", *Journal of Optimization Theory and Applications* 20 (1976) 315–328.

Stewart, G. W., *Introduction to Matrix Computations*, Academic Press (1973)

Todd, M. J., "Quasi-Newton Updates in Abstract Vector Spaces", SIAM *Review*, Vol. 26 No. 3 (July 1984) 367–377.

Wolfe, P., "Convergence Conditions for Ascent Methods", SIAM *Review* 11 (1969) 226–235

Wolfe, P., "Convergence Conditions for Ascent Methods II : Some Corrections", SIAM *Review* 13 (1971) 185–188.